

# Canadian Journal of Research

*Issued by THE NATIONAL RESEARCH COUNCIL of CANADA*

VOLUME 2

FEBRUARY, 1930

NUMBER 2

## ON RECTIFICATION AND DETECTION RECTIFICATION AND POWER PACK DESIGN<sup>1</sup>

BY RICHARD RUEDY<sup>2</sup>

### Abstract

The theory of the mercury arc rectifier is applied to the problem of rectification and detection of alternating currents by vacuum tubes, in radio work, and the meaning of the conventional detector formula is discussed in terms of this theory. For comparison the rectifying action produced in discharge tubes by virtue of their falling characteristic is calculated. The design of filters for smoothing out the rectified current is treated by making use of the results obtained from telephonic transmission lines. By means of the formulæ given for the different cases a comparison can be made between the filters with capacitance input used since 1925, and the type of filters which should be used in order to prevent temporary overloading of the power tubes. The chief experimental methods for testing the theory, and for examining the rectified current, are indicated; the advantages of the glow discharge tube oscillograph are pointed out, and its use in testing work is proposed in place of the methods hitherto followed.

### Introduction

The rapid progress in the broadcasting field did not always allow enough time for careful planning, or for examining in detail the theoretical side of the changes and improvements which have been introduced at various occasions. Cut and try methods play an important part. Thus the filtering devices used in connexion with A.C. supply introduced as early as 1925, and recommended in the instruction sheets of manufacturers of tubes, are known to overload the rectifying tube during the peak value, and to reduce the life of the tube. The problem of the detector is another example in point; the experimental results claimed by different investigators are contradictory, for instance as far as the respective merits of plate and grid leak detection are concerned. On the other hand, the theory of the mercury arc power detector has attained a high degree of development, and it seems worth while to apply it to a series of problems arising in radio reception, in order to gain a sound basis for designing rectifier and detector circuits. From the scientific point of view these devices form a most interesting example of energy transformation.

<sup>1</sup> Manuscript received January 2, 1930.

<sup>2</sup> Testing Department, Silver Radio Factory, Toronto.

# 1. The Problem of rectification and detection

In radio broadcasting, and in the reception of speech and music, the necessity arises of supplying the plates of the tubes with perfectly constant D.C. potential, when 25 to 60-cycle A.C. is the only available source (rectification), and of detecting the incoming signal, that is, of rectifying the incoming high frequency alternations without suppressing the much slower audio-frequency variations. Both tasks are accomplished by means of high vacuum tubes, and in some cases by means of discharge tubes or three-electrode tubes containing inert gases. In the case of the power supply constant voltages of several hundred volts at 25 to 60 cycles from a transformer have to be rectified. In the case of detection the radio frequency  $f_r$ , amounts to 1,000 and more kilocycles, the voltages being as low as a few microvolts or as high as 10 to 20 volts (power detection); this voltage varies from one oscillation to the other according to the law (modulated high frequency wave):

$$\begin{aligned}\bar{e} &= E_r \sin p_r t (1 + m \sin p_a t) \\ &= E_r \left[ \underset{\text{(carrier)}}{\sin p_r t} + \underset{\text{(lower)}}{\frac{1}{2} m \sin (p_r - p_a) t} - \underset{\text{(upper)}}{\frac{1}{2} m \cos (p_r + p_a) t} \right] \underset{\text{(sideband)}}{\end{aligned}$$

$m$  being the degree of modulation,  $p$  equal to  $2\pi$  times the frequency  $f$ . The modulated wave can be considered, mathematically and physically, as the sum of three high frequency waves of constant amplitude. For the human voice,  $f_a$  varies from 60 (bass) to 1300 (soprano), for the piano from 27 to 4138, for the organ from 16 to 4138, hertz apart from overtones which must be present in each source. In the case of speech and music, a few instruments like the stopped organ-pipe, the tuning fork etc. excepted, a number of  $f_a$  are emitted at one and the same time producing the corresponding number of side bands.

In cases where the A.C. voltage is applied to the rectifier by means of a transformer the secondary of which is provided with a centre tap, both halves of a wave can be rectified separately by having each half of the transformer connected to a separate rectifier, or to a two-anode rectifier (full wave or biphase rectification).

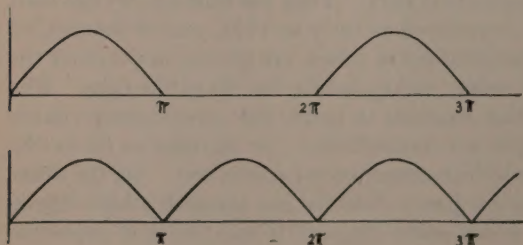


FIG. 1. Rectification of sine wave by ideal rectifier  
(a) single wave (b) full or biphase rectification.

without attenuation, the rectified sine wave  $\sin \theta$  (Fig. 1) is proportional to the Fourier series

In any case the D.C. power obtained at the terminals of a rectifier is hatched or at least pulsating, the current, for instance, consisting of a number of loops in which it drops to zero value every half-cycle. To take the most favorable cases, those where one-half wave is completely suppressed, the remaining half-wave is transmitted



$$\begin{aligned} 1/\pi + \frac{1}{2} \sin \theta - \frac{2}{1.3\pi} \cos 2\theta - \frac{2}{3.5\pi} \cos 4\theta + \dots \\ = 0.32 + 0.5 \sin \theta - 0.22 \cos 2\theta + \dots \end{aligned}$$

for single phase rectification, and to

$$2/\pi - \frac{4}{1.3\pi} \cos 2\theta - \frac{4}{3.5\pi} \cos 4\theta = 0.64 - 0.44 \cos 2\theta + \dots$$

for biphas rectification; that is, the rectifier or detector appears to the output circuit as a generator of a large number of frequencies, and of a certain amount of D.C. A second rectifier would, of course, get rid of part of the higher harmonics, but most often one stage only is used, the outside circuit being arranged in such a way that the D.C. component (0.64 for biphas rectification) is filtered from the different A.C. frequencies by means of self-inductance and capacities.

## 2. The Rectifier E, I characteristics

A rectifying device is essentially a portion of a circuit which does not follow Ohm's law, the resistance depending on the contrary upon the current passing through the device. Strictly speaking, any conductor heating up and changing its temperature periodically when traversed by an alternating current is a rectifier, although the rectifying action is small and tends to disappear at the higher frequencies. In the more important cases of discharges, in gases and in vacuum, the following equations, instead of Ohm's law, represent the relation between voltage and current along certain stages of the discharge. In gases

$$I = A/E$$

$$I = B/\sqrt{E}$$

in vacuum

$$I = kE^{3/2} \text{ (theoretically)}$$

$$I = Ce^{E/E_T} \text{ (small or negative voltages)}$$

$$I = GE \text{ (for } E > 0) \text{ and } I = 0 \text{ for } E < 0 \text{ (ideal rectifier)}$$

$$I = DE^2 \text{ (for very small voltages, and with a different constant for voltages up to half the saturation current).}$$

In the general case, a steady potential  $S$  which is maintaining a constant discharge is applied to the rectifying device, and the alternating current is superimposed, the total voltage acting at the terminals of the detector being  $S + \tilde{e}$ . Given an A.C. voltage  $\tilde{e} = E_o \sin pt$  applied to a rectifier with a known characteristic, the question then arises: How much direct (or audio frequency current) can be obtained (per sec.) in a load  $R$  connected with the rectifier? The load  $R$  may vary from zero (short circuit, zero D.C. voltage across the load) to infinity (open circuit, maximum rectified voltage produced at the output load  $R$ ). The value  $I_m$  of the current available at the terminals of the rectifier, in which the potential  $\tilde{e} = E_o \sin pt$  is superposed upon a steady potential  $E = S$ , that is, the value of

$$I_m = \frac{1}{\tau} \int_0^\tau i \, dt$$

$$\text{where } \tau = 2\pi/p$$

can be calculated, mathematically or graphically (Fig. 2) from the  $E, I$  characteristic as deduced from D.C. measurements, provided the rectifier is operating without lag. This value  $I_m$  represents the current which would flow

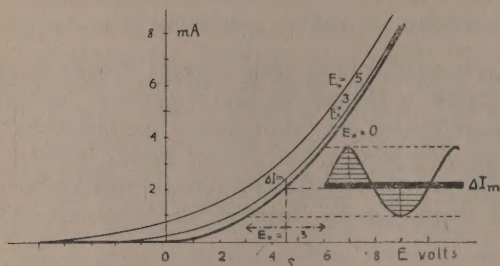


FIG. 2. Graphical analysis of rectifying action when single sine wave is involved (single branched parabola,  $\Delta I$  is the steady D.C. obtainable;  $E_0 = 2$  and  $E_0 = 5$  are computed.)

to measure, the values of  $I_m$  at different values of  $E$ . The curve representing the values of  $I_m$  as a function of  $E$  is called the rectification diagram for a given  $\bar{e}$ . For the particular value  $\bar{e} = 0$ , the rectification diagram coincides with the detector characteristic  $I = f(E)$ . If the rectifier has no lag, the current  $I_m$  produced per cycle is independent of the frequency, so that low frequency current may be used for the study of its action, and the calculations can be carried out for the potential  $E_0 \sin t$  (with  $p = 1$ ) or  $E_0 \sin \theta$  ( $\theta = pt$ ). Outside the rectifying device A.C. and D.C. circulate without influencing each other, at least so long as the circuit contains no parts which are set in motion by external forces. The difference  $\Delta I_m$  between the values of  $I_m$  for  $\bar{e}$  present and  $\bar{e}$  absent is called the detected current. The output circuit to which the rectifier delivers current contains as a rule ohmic resistance  $R$ , and due to the voltage drop across  $R$  the potential  $S$  at the terminals of the detector is smaller than the potential  $B$  of the constant source. An additional more important drop is due to the circulation of the detected current  $\Delta I$ . The steady current  $I$ , and the average current  $I_m$  are given by the intersection of the resistance line (Fig. 3), that is, the line of slope  $\tan a = 1/R$  with the  $I = f(E)$  characteristic, and the rectification diagram, respectively, because at the intersection the current satisfies both equations. For a steady potential of  $S$  volts,

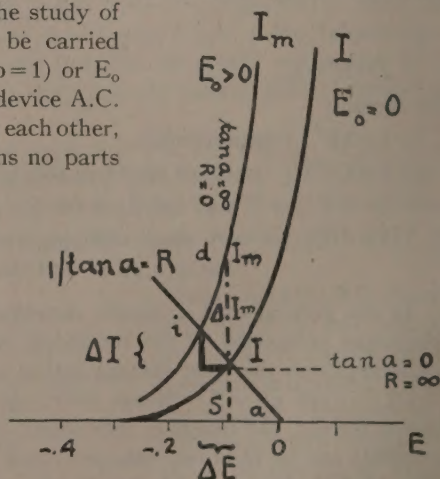


FIG. 3. Rectification diagram and rectification current as a function of  $E = S$  or  $R$  ( $\tan a = 1/R$ ). Approximately an exponential characteristic.



acting at the rectifier terminals, a steady current  $I_i$  is flowing across the detector  $R_d$  and the resistance  $R$  when the potential  $E_o \sin t$  is applied, instead of the current  $I_m$  which would be obtained, in the absence of  $R$ , at the potential  $S$ . Instead of the increase  $\Delta I_m$  taking place on application of the alternating potential in the absence of  $R$  (short circuited detector), the increase  $\Delta I$  is observed when  $R$  is present. To this increase produced by the A.C. source corresponds a voltage drop across  $R$  such that  $\Delta E = R\Delta I$ . The value of  $\Delta E$  is the D.C. voltage produced by the action of the detector upon the superimposed A.C. source. It varies with  $R$ , attaining its greatest value  $\Delta E_\infty$  for  $R = \infty$  (open circuit with  $\tan a = 0$ , a line parallel to the  $E$  axis), and its smallest value 0 for  $R = 0$  (line parallel to  $I$  axis). Designating by  $S_r = 1/R_r$  the slope of the rectification diagram, between points  $i$  and  $d$ , the figure shows that

$$\Delta I = \Delta I_m - S_r \Delta E = \Delta I_m - S_r R \Delta I$$

or

$$\Delta I = \frac{\Delta I_m}{1 + S_r R} = \frac{\Delta I_m}{1 + R/R_r} = \frac{R_r \Delta I}{R + R_r}$$

$$\Delta E = \frac{R \Delta I_m}{1 + S_r R} = \Delta I_m \frac{R R_r}{R + R_r}$$

In order to maintain the operating point  $S$  it is, of course, necessary to adjust the battery voltage  $B$  to the proper value when varying the resistance.

Therefore, as far as the output circuit is concerned, the rectifier acts as a constant emf  $\Delta E_\infty = \Delta I_m / S_r$  having the inner resistance  $1/S_r$  (Fig. 4 and 5), or as a source of the constant current  $\Delta I_\infty$  with the resistance  $1/S_r$  in parallel.

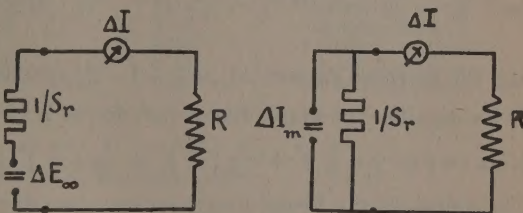


FIG. 4 AND 5. Equivalent rectifier circuit.

For each value of  $E_o$  a different diagram is obtained. The different curves allow to deduce the increase  $\Delta I = I_i - I$  for different values of  $E_o$ , but constant operating potentials  $E$  and a given resistance  $R$ . The curves representing  $\Delta I$  as a function of  $E_o$  constitute the detection diagram of the rectifier for a given  $R$ ; they show immediately whether or not the detected current is proportional to the incoming A.C. potentials.

In connexion with the resistance lines the rectification diagrams for different  $\tilde{e}$  determine therefore completely the properties of a rectifier.

#### RECTIFICATION DIAGRAMS FOR VACUUM TUBES

Except for small or negative  $E$ , the theoretical relation between current and voltage is  $I = kE^{3/2}$ , below the saturation value  $I_s$ . This equation is difficult to handle in the case of rectifiers, and can be replaced in practice by simpler

expressions. With large values of  $\tilde{e} = E_0 \sin t$  no current flows for negative values of  $\tilde{e}$ , that is,  $I = 0$  for  $e < 0$ , whereas for  $e > 0$  the current is represented by a linear function of the voltage:  $I = GE$ . When it is not possible to disregard the curvature of the characteristic, that is, for values of  $E_0$  which are smaller than the potential giving about half the saturation current, the  $3/2$  power curve at the point  $E$  can be replaced, according to Taylor's theorem, by

$$I = f(E+e) = f(E) + \frac{dI}{dE} e + \frac{1}{2} \frac{d^2I}{dE^2} e^2 = I_E + Ge + \frac{T}{2} e^2$$

In many cases a curve  $I = \frac{1}{2} TE^2$  is quite satisfactory ( $T$  being a constant), with  $I = 0$  for  $E < 0$ . For somewhat large distances between filament and anode the simple parabolic law represents in some cases the entire characteristic. Finally, for small and negative voltages the work on the initial velocities of the electrons has shown that the current is given by the exponential function  $I = C e^{E/E_T}$ , where  $E_T$  corresponds to the thermal energy of the electrons at the temperature of the filament.

### 1. Exponential characteristic $I = C e^{E/E_T}$

The total voltage applied to the detector being  $E + E_0 \sin t$  (that is  $S = E$ )

$$I_m = \frac{1}{2\pi} \int_0^{2\pi} i dt = C e^{E/E_T} \times \frac{1}{2\pi} \int_0^{2\pi} e^{\frac{E_0 \sin t}{E_T}} dt = I_{J_0}(jE_0/E_T)$$

and the detected current  $\Delta I_m = I_m - I = I[J_0(jE_0/E_T) - 1]$

$J_0(jx)$  signifies the simple Bessel function of  $jx$  (with  $j = \sqrt{-1}$ )

$$J_0(jx) = 1 + \frac{1}{1!} \left(\frac{x}{2}\right)^2 + \frac{1}{2!^2} \left(\frac{x}{2}\right)^4 + \frac{1}{3!^2} \left(\frac{x}{2}\right)^6 + \dots$$

The values of the Bessel functions may be taken from published tables.

The mean values  $I_m$  of the current flowing under the influence of the A.C. signal are evidently  $J_0(jE_0/E_T)$  times larger for any value of  $E$  than those,  $I$ , obtained without  $E_0 \sin t$ . The new values of the constant current could therefore be obtained from the original  $E, I$  characteristic by a change of the units of the  $I$  axis in the constant ratio  $J_0(jE_0/E_T)$ . Now any shift  $e$  in the operating voltage results also in a change of the units of the  $I$  scale, because  $I = C e^{(S+e)/E_T} = C e^{e/E_T} e^{S/E_T}$ . The curve representing the values  $I_m$  as a function of the variable  $E$  is therefore the exponential characteristic curve shifted with respect to the original  $E, I$  characteristic by the amount  $E_T \ln J_0(jE_0/E_T)$  parallel to the  $E$  axis.

The detection current, namely the increase in current produced by the A.C. signal acting upon the detector, being given by  $\Delta I = I_m - I = I[J_0(jE_0/E_T) - 1]$  is independent of the operating point  $E$ . In other words there is only one detection curve for all possible values of  $E$ . For small values of  $E_0/E_T$  the expression  $J_0(jE_0/E_T)$  reduces to  $(1 + E_0^2/4E_T^2)$ , so that the detection current



$\Delta I = E_o^2/4E_T^2$  is proportional to the square of the amplitude of the incoming signal voltage, whatever the operating point. It is the larger the lower the velocities of the electrons emitted by the detector filament. If the signal voltage is reduced to 0.1 of its value, the detection current is reduced to 0.01 of its strength. The detector is relatively insensitive towards weak signals; it produces a distortion of the amplitudes. The formula is confirmed by the experimental results obtained by different investigators who did not refer to the theoretical basis underlying it. The oxide-coated filament gives a detection definitely superior to that of the tungsten filament coated with thoria, which in turn is better than pure tungsten. For the tungsten filament, indeed,  $E_T = RT = 0.2$  volt at  $T = 2300^\circ \text{C}$ . and for the oxide-coated filament  $E_T = 0.1$  volt at  $T = 1150^\circ \text{C}$ . Measurements on the sensitivity of detector tubes are in agreement with these values (Table I). Moreover, for sufficiently low or negative operating potentials, the sensitivity of a tube remains constant.

TABLE I

VALUES OF  $E_T$  DEDUCED FROM EXPERIMENTAL DATA GIVEN BY F. E. TERMAN (14)

	$M\mu$	$E_T$	$R_r$ ohms
Tube 201A	9	0.23	150,000 and above
171A	3	0.14	200,000 and above
112A	8	0.13	150,000 and above
226	8	0.14	150,000 and above
227	8	0.11	50,000 and above
Screen-grid 222	300	0.20	250,000 and above

The D.C. voltage produced across  $R$  is (page 105)

$$\Delta E = \Delta I_m \frac{RR_r}{R+R_r} = R_r \Delta I_m \frac{R}{R+R_r}$$

$$= \frac{R_r E_o^2}{4E_T^2} \frac{R}{R+R_r} \quad (\text{for small } E \text{ and } E_o)$$

$$\text{or as } \frac{dE}{dI} = R_r = \frac{E_T}{C} \epsilon - E/E_T$$

$$R_r' = \frac{dR_r}{dE} = - \frac{\epsilon - E/E_T}{C}$$

$$\frac{R_r'}{R_r} = - \frac{1}{E_T}$$

$$\text{therefore } \Delta E = - \frac{1}{4} \frac{R_r'}{R_r} E_o^2 \frac{R}{R+R_r}$$

which is the formula for small voltages commonly used in engineering work.

## 2. Parabolic characteristic $I = I_E + f'e + f''e^2$

If the alternating potential  $\tilde{e}$  superimposed at the operating point  $E$  is not large, and  $S$  is in the region where the increase in current with change in  $E$  is only moderate (value of  $E$  corresponding to half the saturation current), the  $3/2$  power law is expressed sufficiently closely by the relation

$$I = kE^{3/2} = f(E + \tilde{e}) = I_E + f'e + f''e^2 + \dots$$

Experiment shows that the coefficients of  $e$  are constant over extended regions of the characteristic. For a 201-A tube, at 20 volts plate potential, L. J.

Peters gives the equation  $\sqrt{I_p} = (4.9e_g + 17.5) \times 10^{-3}$  for  $e_g = -1\frac{1}{2}$  to  $-3$  grid voltage and  $\sqrt{I_p} = (5.75 e_g + 18.5) \times 10^{-3}$  between  $-1$  to  $> +3$  grid voltage. Examining therefore the general expression

$$I = \frac{T}{2} E^2 \text{ with } T = d^2I/dE^2 \text{ constant, for } (E + E_o \sin t)$$

$$I = \frac{T}{2} (E + E_o \sin t)^2 = \frac{T}{2} (E^2 + 2EE_o \sin t + E_o^2 \sin^2 t)$$

$$I_m = \frac{1}{2\pi} \int_0^{2\pi} I dt = \frac{T}{2} E^2 + \frac{T}{2} E_o^2 \text{ (Rectification curve)}$$

$$\Delta I = I_m - I = \frac{T}{4} E_o^2 \text{ (Detection curve)}$$

Thus, as in the case of very small voltages  $E$  and  $\tilde{e}$ , the detected current  $\Delta I$  is for sine waves proportional to the square of the amplitude of the incoming signals. But the constant of proportionality has a different meaning and value, a point very often overlooked.  $T$  varies for different types of tubes from 0.03 to 0.3 mA/volt<sup>2</sup>. With a constant value of  $T$ , the detection current becomes independent of the operating point  $E$ .

The D.C. voltage produced in the process of rectification being

$$\Delta E = R_r \Delta I_m \frac{R}{R + R_r}$$

$$= \frac{R_r}{4} T E_o^2 \frac{R}{R + R_r}$$

or

$$\Delta E = \frac{R_r}{4} \frac{d}{dE} \left( \frac{1}{R_r} \right) \frac{R}{R + R_r}$$

therefore

$$\Delta E = \frac{R_r'}{4R_r} E_o^2 \frac{R}{R + R_r}$$

a formula valid over the whole parabolic characteristic.



For a series of  $E_0 \sin(\theta - \theta_0) + E_1 \sin(2\theta - \theta_1) + \dots$

$$I = \frac{T}{2} (S^2 + E_0^2 \sin^2(\theta + \theta_0) + E_1^2 \sin^2(2\theta - \theta_1) + \dots \\ + 2E_0 E_1 \sin(\theta - \theta_0) \sin(2\theta - \theta_1) + 2E_0 E_2 \sin(\theta - \theta_0) \sin(2\theta - \theta_2) + \dots \\ + 2SE_0 \sin(\theta - \theta_0) + 2SE_1 \sin(2\theta - \theta_1) \\ \int_0^{p/2\pi} I = T/2 \times \frac{1}{2} (E_0^2 + E_1^2 + E_2^2 + \dots).$$

The current from a square law detector is seen to bear no relation to phase displacement; it depends on the maximum amplitude only of the different harmonics. It is proportional to the sum of the square of their amplitudes. On the other hand, the effective or root mean square value of alternating current, the value for which alternating current hot wire or dynamometer instruments are calibrated, is equal to

$$E_{\text{eff}} = \sqrt{E_0^2 + \frac{1}{2}(E_1^2 + E_2^2 + \dots)} = \sqrt{\frac{1}{2\pi} \int_0^{2\pi} I^2 d\theta}$$

As far as alternating currents alone ( $S=0$ ) are concerned, therefore, the detector can be calibrated in terms of R.M.S. values, and adapted to the measurement of alternating currents with D.C. instruments. (The power in the circuit, however, is proportional to  $\frac{1}{2}(E_0^2 + E_1^2 + E_2^2 + \dots)$  only in the case where there is no phase difference between the harmonics, so that for all of the latter the phase displacements between the current and the voltage are the same).

As for a single A.C. component, the detected current is given by  $\Delta I = \frac{1}{4}TE^2$ , the rectifying action does not depend on the operating point; the rectification curve is merely the original curve  $\frac{1}{2}TE_0^2$  displaced parallel to the  $I$  axis, by the amount  $\frac{1}{4}TE_0^2$  for a given  $E_0$ , but only as long as  $S > E_0$ . With sufficient positive potential  $S$  applied to the detector, the rectifying action ought then not to depend on the resistance in the load ( $S > \Delta I R$ ). On the other hand (as shown by Fig. 2), the detected current is the larger the lower  $S$  is for  $S < E_0$ . The voltage produced at open circuit,  $\Delta E_\infty$ , on account of the displacement of the curve, is:

$$\Delta E_\infty = S - \sqrt{S^2 - \Delta I_m T} = S - S\sqrt{1 - E_0^2/2S^2}$$

For values of  $S < E_0$ , that is, if the actual voltage of the detector anode becomes at times negative,  $\Delta E$  increases rapidly with decreasing  $S$ , and becomes theoretically equal to  $E_0$ ; hence a D.C. instrument can be used in series with an extremely high resistance for measuring the maximum voltage amplitude of an alternating current. In order to get an accuracy of 1%, a resistance of 760 megohms would be required.

Like all rectifiers the detector, following a parabolic law or square law, produces new frequencies not present in the original source. Let the incoming

frequency be  $E_o \sin pt$  at the operating point  $S$ ; then, supposing for the moment the more general law  $I = AE + \frac{1}{2}TE^2$ , with  $A=0$  to hold instead of the square law detector hitherto considered, the detector current is

$$\begin{aligned} I &= A(S + E_o \sin pt) + \frac{1}{2}T(S + E_o \sin pt)^2 \\ &= AS + AE_o \sin pt + \frac{1}{2}T[(S^2 + 2SE_o \sin pt + \frac{1}{2}E_o^2(1 - \cos 2pt))] \\ &= AS + \frac{1}{2}TS^2 + \frac{1}{4}TE_o^2 + (A + TS)E_o \sin pt - \frac{1}{4}TE_o^2 \cos 2pt. \end{aligned}$$

The second harmonic is therefore the only one produced with considerable intensity, at least as long as the tube works along the square law curve. A much larger number of frequencies are present, however, when the detector is used in the neighborhood of the origin (for  $E_o > S$ ). For two frequencies arriving at the same time  $a_1 \sin p_1 t + a_2 \sin p_2 t$ , the square law detector gives

$$I = A(a_1 \sin p_1 t + a_2 \sin p_2 t) + \frac{1}{2}T(a_1^2 \sin^2 p_1 t + 2a_1 a_2 \sin p_1 t \sin p_2 t + a_2^2 \sin^2 p_2 t)$$

and, as  $\sin^2 p_1 t = \frac{1}{2}(1 - \cos 2p_1 t)$ ,

$$\begin{aligned} I &= Aa_1 \sin p_1 t + Aa_2 \sin p_2 t + \frac{1}{4}Ta_1^2(1 - \cos 2p_1 t) + \frac{1}{4}Ta_2^2(1 - \cos 2p_2 t) \\ &\quad + \frac{1}{2}Ta_1 a_2 \cos(p_1 - p_2)t - \frac{1}{2}Ta_1 a_2 \cos(p_1 + p_2)t \end{aligned}$$

Summation and difference frequencies are now also present (page 127).

### 3. Straight line characteristic $I = GE$ for $E > 0$ (ideal rectifier)

$$I = 0 \text{ for } E < 0$$

When disregarding the curvature of the  $3/2$  power characteristic in the neighborhood of the origin, as becomes possible for large  $E_o$ , the electron current

may be considered to be determined for positive potentials  $E$  by  $I = GE$ . For the commonly used 280-type tube the value of  $G$ , the inverse of the internal resistance (plate filament resistance) amounts to  $1/250$ , for the grid circuit of 112A tube with +45 volts on the plate to  $1/70,000$  ohms (Fig. 6). Applying a voltage  $\bar{e} = E_o \sin t$  at the terminals of the rectifier, when the operating voltage is  $E = 0$ , gives:

$$I_m = \Delta I_m = \frac{1}{2}\pi \int_0^\pi GE_o \sin t = GE_o/\pi$$

In the general case, a steady potential  $S$  being applied at the detector, the discharge current is given by

$$I = (E_o \sin t - S)/R = \frac{E_o}{R}(\sin t - S/E_o)$$

$$I_m = \frac{E_o}{R} \frac{2}{\pi} \left( \sqrt{1 - S^2/E_o^2} - \frac{S}{E_o} \cos^{-1} \frac{S}{E_o} \right) \quad (\text{Rectification curve})$$

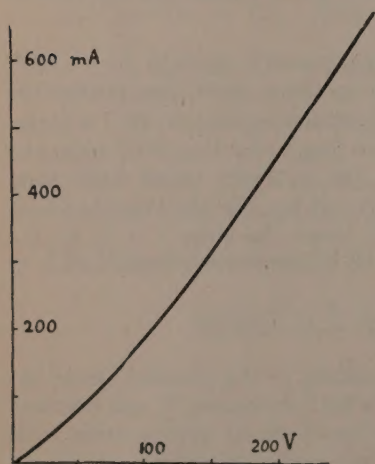


FIG. 6.  $E, I$  characteristic of 280-tube



The complete equation showing the composition of the current is, for full wave rectification

$$I_d = \frac{2}{\pi} \frac{E_o}{R} \left( \sqrt{1 - \left(\frac{S}{E_o}\right)^2} - \frac{S}{E_o} \cos^{-1} \frac{S}{E_o} \right) - \sum_2^{\infty} \frac{2}{n^2 - 1} \frac{\cos n\pi + 1}{\pi} \cdot (\sin a \sin na + \cos a \cos na) \cos n\pi$$

where  $\sin^{-1} S/E_o = a$  or  $\sin a = S/E_o$

$$I_d = \frac{2}{\pi} \frac{E_o}{R} \left[ \left( \cos a - \left( \frac{\pi}{2} - a \right) \sin a \right) + (2/3 \cos^3 a - \cos a) \cos 2t + 1/15 (24 \cos^5 a - 40 \cos^3 a + 15 \cos a) \cos 4t + \dots \right]$$

When at  $S=0$  the steady current  $I_m = \Delta I$  is delivered into a load  $R$ , the voltage drop  $\Delta E$  is produced, and the operating point is thereby shifted by the amount  $R\Delta I = \Delta E$  towards negative values. As far as the current is concerned, this is equivalent to having no load, but a steady opposing e.m.f.  $-\Delta E$  at the rectifier or detector. The values of  $R$  and  $\Delta E$  can therefore be made to correspond to each other (Table II and Fig. 7) and in practical use the load actually determines the operating point. When it is desired to have a tube of the inner resistance  $R_r = 1/G = 1,000$  ohms deliver a direct current of 50 mA per anode, to a load of about 3,000 ohms, corresponding to a voltage drop of 150 volts along  $R$ , Table II gives for 3,000 ohms or about  $2.78/G$ , the value  $\Delta E/E_o = 40/100$ ; therefore 2.5  $\Delta E = 370 = E_o$  volts A.C. have to be applied to the

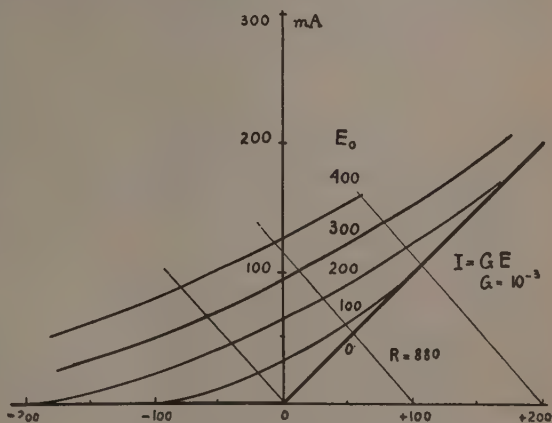


FIG. 7. Straight line characteristic  $I = GE$  for  $E > 0$  and  $I = 0$  for  $E < 0$ , and rectification diagrams for  $E_o = 400, 300, 200$ , and 100 volts.

rectifier. The same result is obtained from the graph: the value 50 mA is reached, for 150 volts operating potential for a point midway between the curve for  $E_o = 300$  and that for  $E_o = 400$ . The maximum potential acting between the two electrodes is evidently equal to  $E_o - E = 220$  volts, producing at that instant a current of 210 mA. The saturation current of the tube has to have at least this value, or else the sine wave would become flattened at the top, when delivering in the neighborhood of 200 mA, and the results of the mathematical analysis of the functioning of the rectifier would no longer apply. It is therefore useful to ascertain not only the value of the direct current drawn, but that of the alternating current circulating in the rectifier as well.

TABLE II  
( $\Delta I$  IN mA)

$\frac{+\Delta E}{E_o}$	-1.0	-0.8	-0.6	-0.4	-0.2	0	(0.2	0.4	0.6	0.8	1.0)
$\frac{R_r \Delta I}{E_o}$	0	27.9	77.4	145	224.5	318	(430 544 677 827 1000) (and subtract 1000 $\Delta E/R_r$ )				
$\frac{R}{R_r}$	$\infty$	29.6	7.8	2.78	0.894	0					
$I_o$	0	5.15	14.0	25.0	37.2	50.0					
$\frac{100N}{GE_o^2}$	0	2.16	4.62	5.76	4.48	0					
$\frac{100\tilde{N}}{GE_o^2}$	0	2.57	7.0	12.5	18.6	25					

NOTE:— $\Delta I = \Delta I_m$  for  $R = 0$  and  $\Delta E = S$   
 $I_m = \Delta I$  for  $\Delta E < 0$

A decrease in the internal resistance of the rectifier increases the rectified current. Thus, for  $1/G = 350$  ohms, and the same conditions as before, the column with  $R = 7.8$  (as  $7.8 \times 350 \approx 3,000$ ) would be used, giving  $\Delta E/E_o = 0.6$ ;  $E_o = 250$  volts would therefore be largely sufficient for obtaining 50mA. For  $R = \infty$  the value  $\Delta E_\infty = E_o$ .

In the steady state the discharge current  $I = \frac{E_o}{R} (\sin t - \frac{S}{E_o})$ , and as current and voltage are in phase, the power used on the alternating current side of the rectifier in the transformer secondary circuit, for instance, is given approximately, per half wave by

$$\frac{1}{\pi} \int_0^\pi E I dt = \frac{E_o^2}{\pi R} \int_{\arcsin \Delta E/E_o}^{\pi - \arcsin \Delta E/E_o} \sin t (\sin t - \Delta E/E_o) dt = \frac{E_o^2}{\pi R} \left( \cos^{-1} \frac{\Delta E}{E_o} - \frac{\Delta E}{E_o} \sqrt{1 - \left(\frac{\Delta E}{E_o}\right)^2} \right)$$

This is equivalent to a maximum current  $I_o \sin t$  with  $(\sin^{-1} \frac{\Delta E}{E_m} = a)$

$$I_o = \frac{E_o}{\pi R} (\pi/2 - a - \sin a \cos a) = \frac{E_o}{2\pi R} (\pi - 2a - \sin 2a)$$

at the potential  $E_o \sin t$ . In the first example cited above  $I_o = 25E_o \times 10^{-5}$ . When the values of  $I_o$  are calculated and added to the table, the power expended on the alternating current side and on the rectifying side in order to get a direct current  $\Delta I$  can be rapidly computed and compared.



## DISCHARGE TUBE RECTIFIERS CONTAINING GASES

Compared with vacuum tubes, discharge tubes have much more varied  $E, I$  characteristics. For steady operation at a certain stage a series resistance is indispensable. Once the discharge, or arc has struck, the potential between the electrodes decreases with increasing current (falling characteristics  $I = 1/E$  or in some cases  $I = 1/\sqrt{E}$ ), and reaches constant values at stronger currents. As soon as the whole cathode becomes covered with glow, the voltage starts to increase quite rapidly, from a few hundred to 1,000 and more volts. By feeding therefore A.C. to a tube with one small and one very much larger electrode, the current meets little resistance when the large electrode is the cathode, but is almost completely suppressed in the following half-wave. As the electrodes are without direct influence upon each other, a second anode can be added without difficulty, in order to take advantage of the full wave. A minimum operating (and starting) potential is required, before the current starts to flow, from 100 to 200 and more volts, depending on the gas and the electrodes, in contrast with vacuum tubes. This disadvantage is partially compensated by the much larger currents obtainable from the glow discharge, at the same voltages. On the other hand, the sudden increase at the breakdown of the gas causes strong harmonics in the current wave, and consequently a more elaborate design of the filter circuit is indicated. Glow rectifier tubes are available giving up to 300 mA of D.C. at 2,000 to 3,000 volts, as well as types adapted for lower values.

1, Characteristic  $E = A/I$ 

This type of falling characteristic is met with in many types of arc discharges; although not often used for rectification, it is of interest for cases in which direct and alternating current are superimposed in an arc. When again  $S + \tilde{e} = S + E_0 \sin t$  volts are applied,

$$I_m = \frac{1}{2\pi} \int_0^{2\pi} \frac{dt}{S + E_0 \sin t} = \frac{2}{\pi \sqrt{S^2 - E_0^2}} \left[ \tan^{-1} \frac{E_0 + S \tan t/2}{\sqrt{S^2 - E_0^2}} \right]_0^{2\pi} \quad \text{for } S > E_0$$

$$= 1/\sqrt{S^2 - E_0^2}$$

For small  $E_0$  and large  $S$ , the detected current  $\Delta I_m$  is

$$\Delta I_m = \frac{1}{S \sqrt{1 - E_0^2/S^2}} - \frac{1}{S} = \frac{1}{S} \left( \frac{E_0^2}{S^2} + \frac{3}{8} \frac{E_0^4}{S^4} + \dots \right)$$

$\Delta I_m$  disappears for large  $S$ , and reaches its largest values for  $E_0$  close to  $S$ . It is approximately proportional to the square of the amplitude of the signal voltage. These properties become evident when a rough graphical solution of the problem is made.

## 2. Inverse square root $I = 1/\sqrt{E}$ .

For this case

$$\begin{aligned}
 I_m &= \frac{1}{2\pi} \int_0^{2\pi} \frac{dt}{\sqrt{S + E_0 \sin t}} = \frac{1}{2\pi} \int_0^{2\pi} \frac{dt}{\sqrt{S + E_0 \cos(\pi/2 - t)}} \\
 &= \frac{1}{2\pi} \int_0^{2\pi} \frac{dt}{\sqrt{S + E_0 - 2E_0 \sin^2(\pi/4 - t/2)}} \\
 &= \frac{1}{\pi \sqrt{S + E_0}} \int_{-\pi/4}^{+3\pi/4} \frac{d\theta}{\sqrt{1 - 2E_0 \sin 2\theta / (S + E_0)}} \\
 &= \frac{1}{\pi \sqrt{E_0} \sqrt{1 + S/E_0}} \int_{-\pi/4}^{3\pi/4} \frac{d\theta}{\sqrt{1 - k^2 \sin^2 \theta}}
 \end{aligned}$$

The values of the elliptical integral may be found in tables, usually up to the upper limit  $\pi/2$  of the integral (values  $K(k)$ ); beyond this limit the integral is given with sufficient approximation by means of the formula

$$F(3\pi/4, k) = 1.5 K(k) + 0.125 k^2 + 0.094 k^4 + 0.07 k^6 + 0.06 k^8 + \dots$$

The result is

$$\begin{aligned}
 S/E_0 &= 65.8, 28.4, 7.33, 3.0, 1.66, 1.3 \\
 I_m \sqrt{E_0} &= 0.03, 0.06, 0.39, 0.62, 0.88, 1.1
 \end{aligned}$$

The graph constructed for  $E_0 = 4$  shows the rectifying effect to be expected (Fig. 8).

### 3. Small cathode discharge tube rectifier

A voltage  $S + E_0 \sin t$  being applied, the current starts suddenly when the potential has reached the value  $A = E_0 \sin t$ ; during the discharge the potential at the electrodes remains at the constant value  $N$ , the normal cathode drop, and the current disappears suddenly at the moment when the instantaneous value of the applied alternating voltage

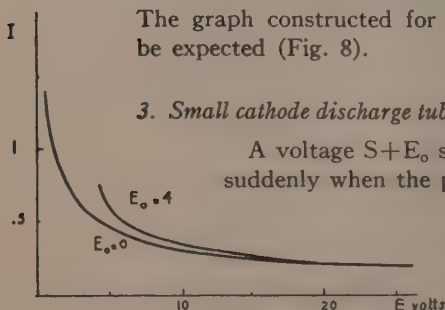


FIG. 8. Characteristic  $I = 1/\sqrt{E}$  and rectification diagram for  $E_0 = 4$  volts.

wave has dropped to  $Z < N$ . The starting and operating potentials differ usually by less than 10 or 20 volts, and can therefore be considered as identical for the study of the rectifying properties so that with an external resistance  $R$ , and  $S = 0$ .



$$I = \frac{E - N}{R} = \frac{1}{R} \left( E_o \sin t - \frac{N}{E_o} \right)$$

Mathematically, although not in the real physical significance, this equation is identical with that of the straight line rectifier. The average current, which is in this case the rectified current, is

$$\Delta I = I_m = \frac{E_o}{R} \frac{\cos A}{\pi}$$

and the equation of the rectified current is for full wave rectification

$$I_m = \Delta I = \frac{2E_o \cos A}{R \pi} - \frac{2E_o}{R} \sum \frac{\cos A \cos nA + n \sin A \sin nA}{\pi} (1 + \cos n \pi) \cos n t.$$

### 3. Filtering the Current Obtained from the Tube

The rectification and detection diagrams furnish the maximum D.C. that can be obtained from a given rectifier. The current as drawn from the terminals of the rectifier is still in reality (Fig. 1) a mixture of D.C. and of A.C., the original input frequency being usually present together with higher harmonics. It becomes therefore necessary to separate the larger D.C. component (64% in bi-phase rectification) from the A.C. components, to filter the output current and to smooth out the pulsations. This can be done in various ways.

(a) In the first place, a by-pass can be provided in the form of a large condenser shunting the output load  $R$  and offering as little resistance as possible to the lowest A.C. frequencies, whilst blocking completely the D.C. component, which finds itself reduced to pass through  $R$ . For full wave rectification the lowest frequency to be considered equals twice the supply frequency (page 102), so that for 25-cycle supply a  $4\mu f$  condenser offers to the A.C. component a resistance equivalent to 800 ohms; roughly, with a 5,000-ohm load only about one-sixth of the lowest frequency would remain in the load. A disadvantage of the method of filtering by a large condenser is that, the condenser representing a low resistance for A.C., large and rapidly varying currents are favored in the rectifier which are capable of overloading the tube as well as the condenser. Where this danger exists, and the voltages concerned are not unduly high, it is preferable to use self-inductance  $L$  to smooth out the current.

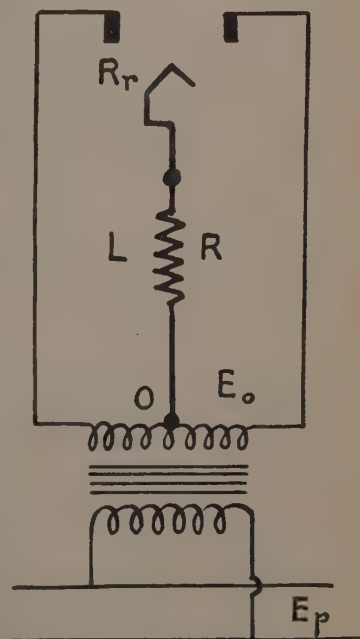


FIG. 9. Rectifier circuit; full wave rectification.

An inductance  $L$  offers a high resistance, equivalent to  $Lp$ , or 314 ohms per henry, to the usually lowest A.C. component, or 6,280 ohms for a 20-henry choke coil; moreover it causes the different A.C. components to lag with respect to each other so that the drain on the rectifier tube becomes more uniform. The exact shape of the rectified current can be calculated for larger A.C. voltages, for which the rectifier is to be assimilated to a constant one-sided resistance  $R$ , the current being suppressed for negative potentials. Considering biphasic rectification, (Fig. 9) the equations for the instantaneous values for the rectified current are

$$\frac{E_o}{R} \left( \sin t - \frac{S}{E_o} \right) = I_m + \frac{L}{R} \frac{dI_m}{dt} = I_m + \tan l \, dI_m/dt$$

The solution of this equation is ( $s = \frac{S}{E_o}$ )

$$I_m = \frac{E_o}{R} \left\{ \cos l \sin (t-l) + \left[ \cos l (\sqrt{1-s^2} \sin l - s \cos l) + s \right] \epsilon^{-t \cot l (1-\sin^{-1} s)} - s \right\}$$

for the first half period for which  $\sin^{-1} \frac{S}{E_o} < t \leq \pi$ . The exponential term causes the smoothing out of the current peaks, and the spreading of the current into the following half-wave of potential. Similar equations are valid for the following half-periods, the value for  $I_m$  at the beginning being equal each time to that at the end of the preceding half-wave, so that finally in the steady state

$$\Delta I = I_m = \frac{E_o}{R} \left( \cos a \sin (t-a) - \frac{\sin 2a}{1 - \epsilon^{-\pi \tan a}} \epsilon^{-t \cot a} - \frac{S}{E_o} \right)$$

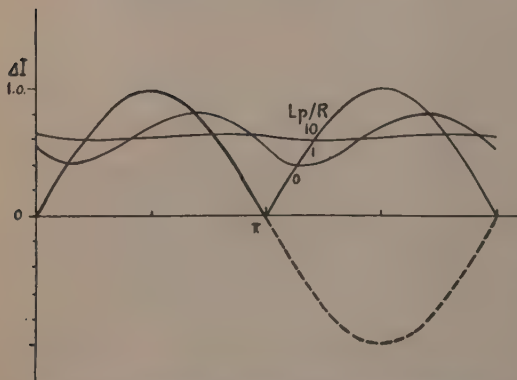


FIG. 10. Filtering by means of choke or inductance  $L$ , for different values of  $Lp/R$ . The action of  $L$  is seen to smooth out the current and to shift the peak values.

To show the wave form of the rectified current, the expression has been evaluated assuming  $E_o/R=1$  for several values of  $\tan l = Lp/R$  ( $\tan l = 0, l = 0$ ;  $\tan l = 1, l = \pi/4$ ;  $\tan l = 10, l = 84^\circ 20'$ ). The term  $S/E_o$  was omitted since it corresponds to taking away a constant quantity  $S/R$ . In the commonly used rectifier tubes of type 280 the internal resistance amounts to 250 to 300 ohms, so that when  $E_o=300$  volts are applied, the ratio  $E_o/R_r$  is equal to unity, and Fig. 10 gives directly the shape of the current wave pro-

duced by this detector and  $L$ . For 50-cycle current,  $\tan a$  is then equal to  $L$ . With loads of the order of 5,000 ohms, the tube resistance may be neglected, and the value of the load  $R$  alone determines the value of  $\tan a = pL/R$ .



The expression for the rectified current developed according to a Fourier series gives for the first terms

$$I = I_m - k \sin (2\theta + b) = 2E_o / \pi R - k \sin (2\theta + b)$$

with  $k = 2E_o / 3\pi pL = 0.212E_o / pL$  and  $\tan b = R / pL$ .

In other words, the rectified current is to be considered as a constant current upon which is superposed an alternating current of amplitude  $0.21E_o / pL$ , with twice the frequency of the original supply. The constant term predominates the more, the lower  $R$  is and the larger  $L$  (Fig. 10).

(b) When  $R$  and  $L$  are shunted by a condenser  $C$ , the detector resistance  $R$ , being neglected, the amount of rectified current decreases. Indeed (Fig. 11)

$$E_o \sin \theta = 1/pC \int i_1 d\theta$$

$$E \sin \theta = Ri_2 + pL di_2/d\theta$$

$$i_1 = E_o \tan m \cos \theta$$

$$i_2 = \frac{E_o}{R} \left[ \cos l \sin (\theta - l) + C e^{-\theta \cot l} \right]$$

$$\tan m = pRC$$

$C$  is determined by the condition that  $I$  returns to the same value after an interval  $\pi$ , so that

$$I = \frac{E_o}{R} \left[ \cos l \sin (\theta - l) + \frac{\sin 2l - \tan m}{1 - e^{-\pi \cot l}} e^{-\theta \cot l} + \tan m \cos \theta \right]$$

$$\Delta I = I_m = \frac{E_o}{R} (2/\pi + S/E_o - \tan l \tan m),$$

and it appears possible to reduce the rectified current to zero. Inductances used in rectifier circuits ought therefore to have as little distributed capacity as possible.

The performance can be improved by introducing a self-inductance  $L$  of the small resistance  $W$  between the rectifier and the condenser, thus doing away with the capacitive input (Fig. 11). The equation for this case can be completely solved (6). After applying the A.C. voltage, there is a period of transients which, however, die down exponentially. The alternating current in the resistance is not far from being inversely proportional to the square of the frequency, but for large values of  $R$ , ( $10^3$  to  $10^4$  ohms) as are used in practice, the decrease is slower, and the combination cannot be said to make a good filter. Moreover, for

$$C = L_1 / (R^2 + p^2 L_1^2) + L_2 / (W^2 + p^2 L_2^2)$$

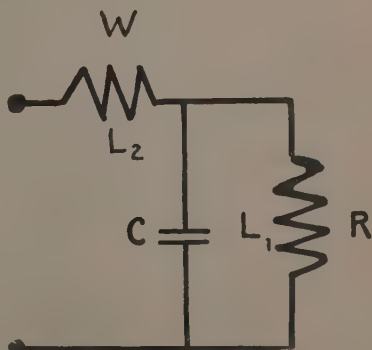


FIG. 11. Filter circuit with combined inductance  $L$  and condenser  $C$ .

current and voltage reach maximum values. If the two  $L$  values are of the same order, and  $W=R$ , so that  $C=2L/(R^2+p^2L^2)$ , then

$$C(R^2+p^2L^2)=2L$$

$$\text{and} \quad p^2 = \frac{2L-CR^2}{CL^2}$$

By taking therefore  $2L \approx CR^2$  the resonance point is made to fall below the lowest frequency to be found in the rectifier output.

All the filtering arrangements described thus far lack sufficient increase in attenuation when higher frequencies i.e., those for which the ear is most sensitive, are considered.

The addition of a second condenser  $C_1$ , shunting the rectifier output terminals, creates a group which is used in the power pack design for simple A.C. sets, and which incidentally is also characteristic of the actual circuit in amplifier tubes. The equations which are valid for the current from the moment when the voltage is applied until the steady state is reached can be solved by graphical or numerical integration.  $E \sin \theta$  being the potential applied to the first condenser, the voltage across  $C$  is equal to  $kE_o \sin(\theta-1)$  with

$$k = \left/ \frac{Z_2}{Z_1+Z_2} \right/ , \quad L_1=0, \quad L_2=L, \quad W=0 \text{ (Fig. 11)}$$

$$E_o \sin \theta = 1/p \int C i_1 d\theta$$

$$E_o \sin \theta = pL di_2/d\theta + kE_o \sin(\theta-1)$$

$$Z_1 = pL$$

$$Z_2 = 1/(pC_2 + 1/R)$$

$$Z_2/(Z_1+Z_2) = R/(p^2RLC_2 + pL + R) = k$$

$$kE_o \sin(\theta-1) = Ri_4 = R(i_2 - i_3)$$

$$= 1/pC_2 \int i_3 d\theta$$

therefore

$$pkE_o C_2 \cos(\theta-1) = i_3$$

$$pkC_2 E_o \sqrt{1 - \sin^2(\theta-1)} = i_3$$

$$kE_o \sin(\theta-1) = R(i_2 - pkC_2 E_o \sqrt{1 - \sin^2(\theta-1)})$$

Putting  $kE_o \sin(\theta-1) = z$

$$0 = z^2(1 + p^2 C_2^2 R^2) - 2Ri_2 z + (R^2 i_2^2 - p^2 k^2 C_2^2 R^2 E_o^2)$$

$$Z = kE_o \sin(\theta-1) = \frac{Ri_2 \pm \sqrt{R^2 i_2^2 - (1 + p^2 C_2^2 R^2)(R^2 i_2^2 - k^2 p^2 C_2^2 R^2 E_o^2)}}{1 + p^2 C_2^2 R^2}$$

$$= \frac{Ri_2 \pm pC_2 R \sqrt{k^2 E_o^2 (1 + p^2 C_2^2 R^2) - R^2 i_2^2}}{1 + p^2 C_2^2 R^2}$$

$$= ai_2 \pm B\sqrt{C - R^2 i_2^2} = ai_2 \pm b\sqrt{c^2 - i_2^2}$$

One of the two values for  $kE_0 \sin (\theta - 1)$  has to be introduced into the equation for the circuit containing  $L$

$$0 = ai_2 \pm b\sqrt{c^2 - y^2} + pL di_2/d\theta - E_0 \sin \theta$$

$$a = R/(1 + p^2 C_2^2 R^2)$$

$$b = pC_2 R^2/(1 + p^2 C_2^2 R^2) \quad c^2 = k^2 E_0^2 (1 + p^2 C_2^2 R^2)/R^2$$

The differential equation can be solved by well-known approximation methods (12). For such values of  $C, R, L$  as occur in practical work:

$$R \sim 1,000 \text{ ohms}$$

$$p = 3 \times 10^3$$

$$1 + p^2 C^2 R^2 \approx 1 \text{ to } 2$$

$$C \approx C_1 \approx 3 \mu f$$

$$L = 10$$

$$a \approx 1,000$$

$$E_0 = 1,000 \text{ volts}$$

$$pL = 3.10^3$$

$$b \approx 1,000$$

$$k = 0.7$$

$$c \approx 1.81 k^2 \text{ to } 2 k^2$$

$$c^2 \sim 1$$

The solution of the equation, taking the positive root, is represented in Fig. 12, and shows a case where the steady state is reached rapidly. (For the equations applying to the steady state, see also the following section.)

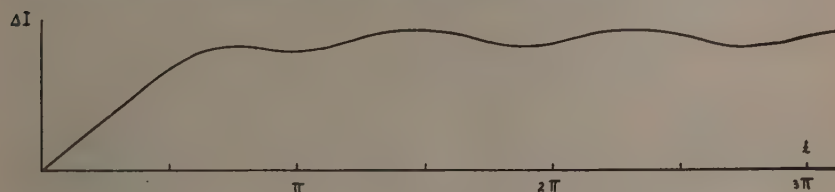


FIG. 12. Rectified current as a function of time when  $E_0$  is suddenly applied.

(c) A third method of smoothing out the current takes advantage of a difficulty met with in telephonic transmission. At the end of long transmission lines the speech can become unintelligible owing to distortion, that is, the different harmonics suffer a different degree of attenuation; generally the weakening effect tends to increase with the frequency. The higher of the two characteristic frequencies of the vowel *ee*, in three ( $f \approx$  about 3100), for instance, is not well transmitted, so that *ee* sounds somewhat like *oo* in two. If  $E_s$  volts are applied to a practically endless line, the voltage at the distance  $x$  will be  $E = E_s e^{-\gamma x}$  where the attenuation constant  $\beta$  is determined by

$$\gamma = \beta + j\alpha = \sqrt{jpC(R + jpL)}$$

$$\beta = \sqrt{\frac{1}{2}pC(\sqrt{R^2 + p^2L^2} - Lp)}$$

The phase difference and wave-length constant

$$\alpha = \sqrt{\frac{1}{2}pC(\sqrt{R^2 + p^2L^2} + Lp)}$$

$C$  being the capacity,  $R$  the resistance,  $L$  the self-inductance referred to one mile length. Thus for a line having  $R = 88$  ohms,  $L = 0.001$  henry,  $C = 0.054$  microfarad, the attenuation constant  $\beta$  becomes 0.106, and after 35 miles an



angular frequency  $p=5000$  suffers a reduction which decreases its amplitude to one fortieth of its initial value. Between  $E_s$  and  $I_s$ , the value of voltage and current at the sending end, and  $E_r$  and  $I_r$ , the values at the receiving end of the line element of length  $l$ , the well known relations hold:

$$E_r = \cosh \gamma l E_s + Z_o \sinh \gamma l I_s$$

$$I_r = \sinh \gamma l E_s / Z_o + \cosh \gamma l I_s$$

$$Z_o = \sqrt{(R + jpL) / jpC}$$

It is possible to replace the line by a shorter artificial line with lumped capacity and impedance (Fig. 13). For a single mesh of this line

$$E_r = E_s [jpC(jpL + R) + 1] - I_s(jpL + R)$$

$$I_r = -E_s [p^2 C^2 (jpL + R) - 2jpC] + I_s [1 + jpC(jpL + R)]$$

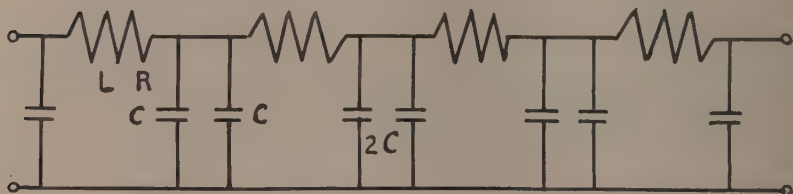


FIG. 13. Low pass filter

Two lines consisting of  $n$  elements, one having distributed impedances, and the other lumped impedances, are equivalent if

$$\cosh \gamma l = 1 + jpC(jpL + R) = \cosh \gamma_1 \text{ (say)}$$

$$Z_o \sinh \gamma l = -(jpL + R)$$

$$\frac{\sinh \gamma l}{Z_o} = p^2 C^2 (jpL + R) - 2jpC$$

Whereas  $\cosh \gamma l$  determines the properties of a uniform line consisting of pieces of length  $l$ , the expression  $1 + jpC(jpL + R) = \cosh \gamma_1$  is decisive for the properties of the artificial line composed of an assemblage of lumped impedances. Neglecting the resistance  $R$ , and putting  $\cosh \gamma_1 = \cosh(\beta_1 + j\alpha_1)$

$$E_r = uE_s - jpLI_s$$

$$I_r = E_s L(1 - u^2) / jp + I_s u$$

where  $u$

$$u = 1 - p^2 LC = (1 - 2p^2 / p_o^2), \text{ with } p_o \text{ corresponding to}$$

the natural frequency of the mesh

$u = \cosh \gamma_1 = \cosh(\beta_1 + j\alpha_1) = \cosh \beta_1 \cos \alpha_1 - j \sinh \beta_1 \sin \alpha_1$  and evidently  $j \sinh \beta_1 \sin \alpha_1 = 0$ ,

that is either  $\beta_1 = 0$ ;  $\cosh \beta_1 = 1$ ,

$\cos \alpha_1 = u$  or  $\cos \alpha_1 = 1 - 2p^2 / p_o^2$  or

$0 < p^2 / p_o^2 \leq 1$ ; no attenuation  $\beta_1 = 0$ ,

or  $\alpha_1 = \pi$ ;  $\cos \alpha_1 = -1$ ;  $\cosh \beta_1 = 2p^2 / p_o^2 - 1$ ;

$p^2 / p_o^2 > +1$  and  $\beta_1 = \cosh^{-1}(2p^2 / p_o^2 - 1)$

so that, for frequencies for which  $p^2/p_o^2 > 1$ , the voltage amplitude  $E_s$  is reduced to  $E_s e^{-\text{arc cosh } u}$ , after the first mesh, and to  $E_s e^{-n \cdot \text{arc cosh } u}$  after the  $n^{\text{th}}$  element.

For  $f_o = 12$  and  $f = 16$  or  $p_o = 75$  and  $p = 100$ ,  $\cosh^{-1} 23/9 = 1.59$ , so that after passing through the first element the voltage amplitude of a 16-cycle component will amount to 0.2, after passing through the second element to 0.04, and after the third to 0.008 of the original value. In order to make  $p = 100$ , the value corresponding to the lowest frequency which the ear perceives as sound, taking  $L = 10$  henry,  $C$  has to be given a value equal to  $\sqrt{2/LC} = 100$  or  $C = 20$  microf.

These formulæ are only valid for an infinite line, but it is known that any shorter line behaves as an endless line when the load  $Z_r$  into which it delivers current is equal to the characteristic impedance  $Z_o$  of the line, that is when for each element, including the generator  $Z_s$ ,

$$Z = Z_o = Z_s$$

$$\text{where } Z_o = \sqrt{\frac{L(1 - Rj/pL)}{1 - p^2LC/2 + (jpCR/2)2C}} = \sqrt{\frac{L}{2C}} \sqrt{\frac{1 - jR/pL}{1 - p^2LC/2 + jpCR/2}}$$

$$= \sqrt{\frac{L}{2C}} \sqrt{\frac{1 - p^2LC/2 - CR^2/2L - jR/pL}{(1 - p^2LC/2)^2 + p^2C^2R^2/4}}$$

For values a little above the resonance frequency  $Z_o = \infty$ , and the filter action becomes evident. When the coil resistance is small, the load resistance would have to have a value of about  $\sqrt{10^7/40} \approx 10^3$  ohms for  $f = 16$ .

It is often advantageous to suppress the half-size condensers  $C$  which are placed at the beginning and at the end of the filter, but to start (from mesh zero), and to end (into mesh  $n$ ) with an inductance  $L/2$ . Calling  $Z_s$  the impedance of the source, the current flowing in the first inductance of value  $L$  (mesh 1) is  $E_s e^{-\beta_1/(Z_s + Z_o)}$  where  $\beta_1$  is again equal to  $\cosh(2p^2/p_o^2 - 1)$  for a filter without resistance; the current traversing the third coil (mesh 2) is  $E_s e^{-2\beta_1/(Z_s + Z_o)}$  etc. The value of  $Z_o$  is equal to the total impedance of the filter chain

$$Z_o = \sqrt{\frac{L}{2C}} \sqrt{1 + R^2C/2L - LCp^2/2 + j(pCR - R/pL)}$$

When the characteristic impedance of the line is equal to that of the source from which it receives, and also equal to that of the load into which it delivers, current, this type of filter consisting of a finite number of elements behaves again as an assembly of an unlimited number of elements. The filter with the inductive element at beginning and end draws a less heavy peak current than the filter with the capacitance input, and is less likely to produce excessive peak voltages at the rectifier tube, owing to the fact that the different harmonics show a lag which increases with  $p$  (formula for  $Z_o$ ).

Input, output and filter impedance depend in general upon the frequency, so that they match only for a definite value of  $p$ . For other frequencies, the equations have to be solved from element to element (Fig. 14). The general

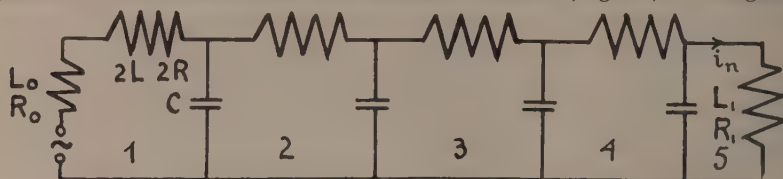


FIG. 14. Low pass filter with input and output circuit.

solution for the steady state is of the form expressed by the equation,  $I_m$  being the current entering the  $n^{\text{th}}$  mesh (3)  $I_m = \sin \xi \cos 2(n-m) \xi f_1(t) + \sin 2(n-m) \xi f_2(t)$ , an equation which is satisfied on condition that

$$-4 \sin^2 \xi = h = 2jpC(R + jpL) \text{ or } 2 \cos \xi = \sqrt{h+4}$$

$$h_1 = jpC(R_1 + jpL_1)$$

$$h_o = jpC(R_o + jpL_o)$$

$$f_1(t) = jpCE_0 e^{jpt} / N = jpCE / N$$

$$f_2(t) = (h_1 - 1 + \cos 2\xi) jpCE / N$$

$$N = h_o h_1 \sin 2(n-1) \xi - 4 \sin^2 \xi \sin 2n \xi + 2(h_o + h_1) \sin \xi \cos(2n+1) \xi$$

so that the current circulating in the last element ( $n=m$ ) becomes

$$I_n = jpCE \sin 2\xi / N.$$

The natural frequencies of the filter chain are given by  $N=0$ .

For  $n=2$

$$I_2 = \frac{jCpE}{h + (h+1)(h+h_o+h_1) + h_o h_1}$$

for  $n=3$

$$I_3 = \frac{2jpCE}{(h+h_o+h_1)(h+1)^2 + (h+1)(h_o h_1 + 2h) + h h_o + h h_1 + h_o h_1}$$

and for  $n=4$

$$I_4 = \frac{2jpCE}{(h^4 + 6h^3 - 10h^2 - 76h + h^2 h_o h_1 + 4h h_o h_1 + 3h_o h_1) + (h_o + h_1)(5h^2 + 6h + 1)}$$

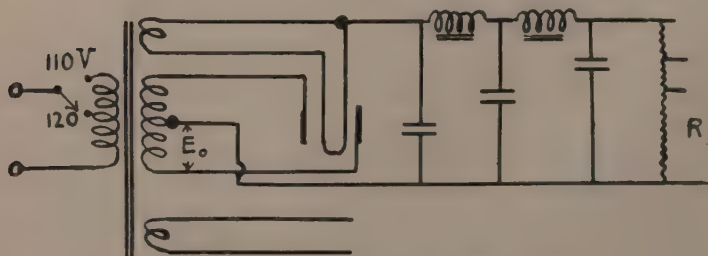


FIG. 15. Filter circuit as used on A.C. sets.



A drawback common to all the filter-circuits described and in use (Fig. 15) is that the main current passes through all the choke coils, so that the inductance varies with the load, the strength of the harmonics depends on the current drawn, and the hum is particularly strong during the silent periods of a broadcast. It would be advantageous to by-pass the A.C. components through a set of parallel filter-lines,  $L_1, C_1, L_2, C_2$ , etc., each designed to tune out one particular harmonic (Fig. 16); this requires, of course, an exact knowledge of the current wave.

Finally, although the considerations have been apparently restricted to rectifiers with two electrodes, (power rectifier, grid detector), they apply to three-electrode tubes as well. This is because grid and plate can be replaced, as far as the emission current is concerned, by a single electrode, of potential  $\mu E_g + E_p$ , and occupying the position of the grid.

In order to allow a comparison between the two different types of filters discussed in a previous section, the attenuation constant  $\beta$  has been calculated (a) for a filter with capacitance input, consisting of a series of inductances  $L$ , with condensers  $C$  in parallel

$$L = 9 \text{ henry} \quad R = 100 \text{ ohms} \quad C = 2 \text{ mcf}$$

$$\cosh \gamma = \cosh(\beta + j\alpha) = b + ja = (1 - p^2 LC) + jpRC$$

$$\sinh^2 \beta = -\frac{1}{2}(1 - b^2 - a^2) + \sqrt{a^2 + \frac{1}{4}(1 - b^2 - a^2)}$$

(b) for a filter in which the first and last components are an inductance  $L/2$ .

$$L = 10 \text{ henry} \quad R = 100 \text{ ohms} \quad C = 2 \text{ mcf.}$$

$$\beta = \sinh^{-1} \sqrt{\frac{p^2 L C}{2}} \sqrt{\sqrt{\frac{R^2}{p^2 L^2} + \left[1 - \frac{p^2 L C}{4} \left(1 + \frac{R^2}{p^2 L^2}\right)\right]^2} - \left(1 - \frac{p^2 L C}{4} \left(1 + \frac{R^2}{p^2 L^2}\right)\right)}$$

The maximum tolerable hum in a loud speaker corresponds to about five volts for a 60-cycle component, and 0.5 volt for a 120-cycle component. The formulæ for  $\beta$  show that if condensers of double size are used in case b, the attenuation is the same in both cases.

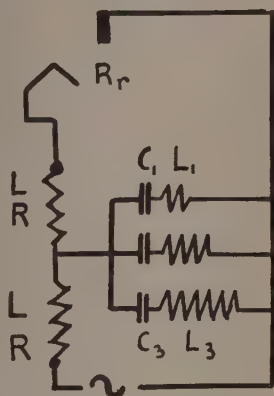


FIG. 16. Filter elements in parallel.

TABLE III  
 ATTENUATION CONSTANT (CASE A)

p	$p^2LC$	pRC	$1-b^2-a^2$	$\sinh \beta$	$\beta$	$\alpha$
00	0.18	0.02	0.33	0	0	24°
200	0.72	0.04	0.92	0	0	38°50'
300	1.62	0.06	0.62	0	0	33°50'
400	2.88	0.08	-2.54	1.59	1.25	
500	4.50	0.10	-11.26	3.36	1.93	
1,000	18	0.2	-288	16.97	3.54	
2,000	72	0.4	-5040	71	4.96	
4,000	288	0.8	-82370	287	6.3	

 TABLE IV  
 ATTENUATION CONSTANT (CASE B)

p	$R^2/p^2L^2$	$p^2LC/4$	$\sqrt{p^2LC/2}$	$\sinh \beta$	$\beta$
100	0.01	0.05	0.32	0	
200	0.0025	0.2	0.63	0	
300	0.0009	0.45	0.95	0	
400	0.00063	0.80	1.27	0	
500	0.0004	1.25	1.58	1.12	0.96
1,000	0.0001	5.0	3.16	9.94	3.0
2,000	0.00003	20	6.33	39.0	4.4
4,000		80	12.65	158.9	5.8
8,000		320	25.30	644.4	7

#### 4. The Measurement of Overtones Produced by the Rectifier (Hum, ripple)

In the more complicated networks used for filtering the rectifier current, the calculations for deducing the exact shape of the current are sometimes cumbersome, and experimental methods are often more expedient. This is

more truly so as the human ear is unequally sensitive to sounds of different frequencies, but identical energy; a sound of frequency of 1,600 hertz appears to the ear as 560,000 times stronger than a sound of 100 hertz involving the same physical energy falling upon the ear. The unassisted ear seems to be most sensitive to frequencies of towards 3,000 hertz, intensities of as low as  $3 \times 10^{-6}$  micro-ergs per sq. cm. per sec. being noticeable at that pitch. It is therefore possible for a high overtone component to have a very small amplitude, but to affect the ear nevertheless more than a lower tone.

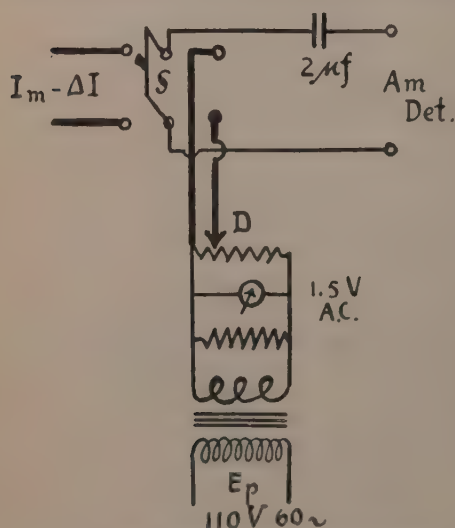


FIG. 17. Circuit for hum measurement.

A rough, but commonly used, method to study the evenness of the rectified current consists in comparing the ripple or hum it produces, as for instance in a telephone receiver or loud-speaker, with the ripple produced by a weak 60-cycle current obtained from the mains by means of a step-down transformer working into a calibrated ohmic resistance,  $R$  (Fig. 17). A double-pole double-throw switch allows part of the rectified current (or the loud-speaker coil), to be applied to grid and filament of a three to four-stage resistance coupled amplifier containing a volume control (variable resistance). The A.C. potential produced in the last stage is measured by a vacuum tube voltmeter, in which it causes a certain current to flow. By throwing the switch to the other side, the 60-cycle current is introduced, its potential being varied at  $D$  until the same reading is obtained as previously with the power pack. The amplifier must of course be so constructed that it transmits readily the lowest frequencies from 20 to at least 100 cycles per sec. When the ripple produced by these low frequencies has a small value, the higher harmonics are as a rule much fainter, so that the method constitutes a test, but only so long as rectifiers of the same design are compared.

The more significant ratio of the amplitude or energy of the overtones as compared to the fundamental can be determined by the well known method of the A.C. bridge, a square law detector being conveniently used as an indicating instrument together with a thermo-couple. If three arms of the bridge are constituted by ohmic resistance, the fourth by ohmic resistance  $R_4$ , capacity  $C$  and self-inductance  $L$  in series, then the bridge will be balanced for the frequency determined by

$$LCp^2 = 1$$

when

$$R_1 = kR_2$$

$$R_3 = kR_4$$

$k$  being chosen in the neighborhood of 0.5 for the sake of good sensitivity. For any other higher frequency a potential drop  $v_h$  exists across the detector amounting to ( $V_h$  being the voltage of the particular frequency in the source, at SS, Fig. 18):

$$\begin{aligned} v_h &= V_h \frac{R_1}{R_1 + R_2} - V_h \frac{R_3}{R_3 + R_4 + jpL + 1/jpC} \\ &= V_h \left( \frac{k}{k+1} - \frac{kR_4}{R_4(k+1) + jpL(h-1/h)} \right) \\ &= V_h \left( \frac{k}{k+1} \frac{1}{1 - j \frac{R_4(k+1)}{pL(h-1/h)}} \right) \end{aligned}$$

Therefore apart from the phase difference

$$v_h = V_h \frac{k}{k+1} \frac{1}{\sqrt{1 + R_4^2 (k+1)^2 / L^2 p^2 (h-1/h)^2}}$$



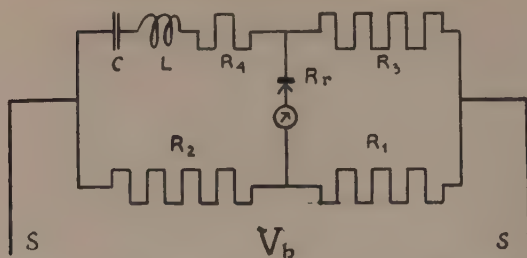


FIG. 18. A.C. bridge for determination of harmonic content.

When  $R_4$  is infinite (open circuit),  $v_h = V_h k / (k+1)$  and  $v_h$  reaches its maximum value, permitting at the same time to deduce  $V_h$  by a simple calculation; if  $k=0.5$ , for instance,  $V_h = 3v_h$ . By taking  $R$  small,  $L$  comparatively large, whilst endeavoring to keep its ohmic resistance low, the value of unity can be realized for the third factor, at least for the high values of  $p$  to which our ear is particularly sensitive.

$$v_h = V_h \left( \frac{k}{k+1} \right) \left[ 1 - \frac{1}{2} \left[ \frac{R_4 (r+1)}{pL(h-1/h)} \right]^2 + \frac{3}{8} \left[ \quad \right]^4 - \dots \right]$$

$$\text{or } v_h = V_h \frac{k}{k+1} \left( 1 - \frac{1}{2} \frac{(r+1)^2 \tan^2 l}{(h-1/h)^2} \dots \right)$$

This voltage acting upon the square law detector produces a direct plate current proportional to

$$I_m = \text{const } V_h^2 \left( \frac{h}{h+1} \right)^2$$

or, for a series of harmonics being present at the same time

$$I_m = \text{const} \left( \frac{h}{h+1} \right)^2 \sum_2^{\infty} V_h^2$$

that is, the current is proportional to the sum of the squares of the amplitudes. When the fourth bridge arm ( $R_4, L, C$ ) is closed and adjusted until minimum reading of the detector current is observed, the strong fundamental must be balanced, and the detector current is proportional to  $\sum_2^{\infty} V_h^2$ . When the arm is now opened, the detector current comprises all the frequencies present and the current is proportional to  $\sum_1^{\infty} V_h^2$ . The two readings therefore give an immediate comparison between the content in higher harmonics and the total A.C. energy, or that of the fundamental whenever the amount of higher harmonics is small. The calibration of the detector meter can be avoided by applying the source, in the case of the open  $R_4$  branch, to a potentiometer ( $R_3, R_6$ ):

$R_6$  is adjusted until the same deflection is obtained as previously with  $R_4$  closed. The ratio  $R_5$  to  $R_6$  gives the ratio of the higher harmonics to the total A.C. content; more accurately the ratio  $R_5$  to  $(R_6 + \text{bridge resistance})$  has to be used, as the bridge arms  $R_1$ ,  $R_2$  are in parallel with  $R_6$ . A simple set up as proposed in the Fig. 19 is very convenient; resistance  $R_7$  can be added, so as to give to the whole bridge a resistance comparable with the load into which the rectifier circuit works.

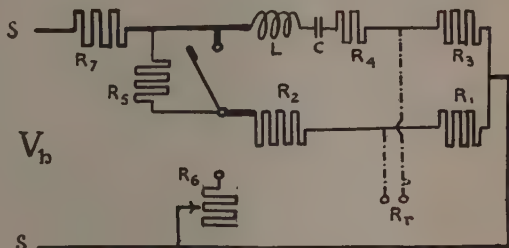


FIG. 19. Circuit for use with A.C. bridge.

The A.C. bridge would in principle allow to determine separately the strength of any particular harmonic; it is only necessary to tune the fourth arm to the frequency, and to read the reduction of the potential across the detector. When however one of the higher harmonics is eliminated, the harmonics in the immediate neighborhood will also be partially balanced, so that corrections have to be calculated for each step.

For gaining detailed information as to the actual frequencies present and their strength, another method is therefore advisable and a difference tone method is used to advantage. A steady pure sine wave, of constant amplitude, but adjustable frequency is added to the unknown frequencies, and the whole current is rectified by a square law detector. The current drawn from the detector plate contains the original frequencies and their harmonics, as well as summation and difference frequencies of the original and the superimposed current waves (page 110). A very low difference note is produced amongst others, whenever the searching frequency happens to coincide approximately with a frequency present in the mixture. These low notes may be isolated by means of a filter having a cut-off above 20 cycles, and may be measured, if necessary after amplification, by a vibration galvanometer tuned to 20 cycles, for instance, when  $f$  is varied from low to higher values the maximum deflections of the galvanometer indicate the presence each time in the mixture of a frequency  $f - f_0 = 20$ . So long as the amplitude of the search frequency is kept constant (heterodyne oscillator), the amplitude of the difference "tone" 20 is proportional to the amplitude of that frequency in the mixture which causes its production; thus the relative, or when the galvanometer is calibrated, the actual amplitudes of the different frequencies can be determined. This method is applicable to a great many acoustical and electrical problems.

Mention must now be made of graphical methods for analyzing a mixture of alternating currents, when an oscillograph is available. This procedure, as is well known from the study of power generator currents, furnishes the frequency, amplitude and phase difference of the components; it may appear therefore to

be too complicated for radio circuits, as the ear is not affected by phase differences. Considering, however, the convenient forms of oscillographs now available, and the high state of development which the mathematical treatment has reached, this type of harmonic analysis is not more difficult to handle than purely electrical methods for deriving the amplitude and frequency of a considerable number of harmonic components. In the case of the fundamental and the first harmonics as produced in the rectification process, representing  $n$  different frequencies, a total of  $2n$  ordinates  $Y_1, Y_2, Y_h, Y_{2n}$ , are measured upon the oscillogram, at intervals of  $\pi/n$ . Assuming the shape of the current or voltage wave to be given by

$$Y = \frac{1}{2}a_0 + a_1 \cos \theta + \dots + a_n \cosh \theta + \dots + a_{n-1} \cos(n-1)\theta + a_n \cos n\theta \\ + b_1 \sin \theta + \dots + b_n \sinh \theta + \dots + b_{n-1} \sin(n-1)\theta$$

it can be shown that

$$\frac{1}{2}a_0 = \frac{1}{2n} \sum_{k=1}^{2n} y_k$$

$$a_1 = \frac{1}{n} \sum_{k=1}^{2n} y_k \cos \frac{\pi k}{n}$$

$$a_2 = \frac{1}{n} \sum_{k=1}^{2n} y_k \cos \frac{2\pi k}{n}$$

$$a_h = \frac{1}{n} \sum_{k=1}^{2n} y_k \cos \frac{h\pi k}{n} \quad ; \quad b_n = \frac{1}{n} \sum_{k=1}^{2n} y_k \sin \frac{h\pi k}{n}$$

( $n$  being constant,  $k$  varying from 0 to  $2h$ ,  $h$  the number of the harmonic). It is necessary to calculate  $1/n \times \sin \pi k/n$ , and  $1/n \times \cos \pi k/n$ , which can be done once for all if the number of harmonics to be considered in the particular rectifier etc., has been decided upon (see for instance Table III for  $n=10$  and  $n=12$ ). Experience and theory have shown that a correction has to be applied to the coefficients thus calculated, in the form of the factor

$$f_n = \left[ \frac{\pi h/2n}{\sin \pi h/2n} \right]^2$$

(error function), with which the results  $a_h$  and  $b_h$  have to be multiplied, when the accuracy of the measurements warrants it.



TABLE V  
CALCULATED VALUES FOR  $1/n \sin \pi k/n$  AND  $1/n \cos \pi k/n$

k	$\frac{1}{10} \sin \frac{k\pi}{10}$	$\frac{1}{10} \cos \frac{k\pi}{10}$	$\frac{1}{12} \sin \frac{k\pi}{12}$	$\frac{1}{12} \cos \frac{k\pi}{12}$	h	$f_{10}$	$f_{12}$
0	.0000	.1000	.0	.0833	0	1.	1.
1	.0309	.0951	.0216	.0805	1	1.01	1.
2	.0588	.0809	.0417	.0722	2	1.03	1.02
3	.0809	.0588	.0581	.0581	3	1.08	1.05
4	.0951	.0309	.0722	.0417	4	1.16	1.10
5	.1000	.0000	.0805	.0216	5	1.24	1.16
6	.0951	-.0309	.0833	.0	6	1.37	1.25
7	.0809	-.0588	.0805	-.0216	7	1.53	1.33
8	.0588	-.0809	.0722	-.0417	8	1.73	1.46
9	.0309	-.0951	.0581	-.0581	9	2.02	1.63
10	.0000	-.1000	.0417	-.0722	10	2.47	1.84
11	-.0309	-.0951	.0216	-.0805	11		2.11
12	-.0588	-.0809	.0	-.0833	12		2.47
13	-.0809	-.0588	-.0216	-.0805			
14	-.0951	-.0309	-.0417	-.0722			
15	-.1000	.0000	-.0581	-.0581			
16	-.0951	.0309	-.0722	-.0417			
17	-.0809	.0588	-.0805	-.0216			
18	-.0588	.0809	-.0833	.0			
19	-.0309	.0951	-.0805	.0216			
20	-.0000	.1000	-.0722	.0417			
21			-.0581	.0581			
22			-.0417	.0722			
23			-.0216	.0805			
24			-.0000	.0833			

The glow discharge oscillograph seems to be particularly adapted to this type of work, on account of the ease with which the current can be handled and recorded. At the normal cathode drop  $N$ , which is considerably lower than the potentials  $E_0$  obtained from the power pack, at least when inert gases are used for filling the tube, the negative glow covers a certain portion „ $l$ ” of the negative electrode and

$$l = L i$$

$L$  being the area of the glow at one mA current. The discharge current  $i$  is

$$i = \frac{E - N}{R}$$

The current density is constant at the given pressure of the gas, and if the voltage of the source varies by  $\Delta E$  the voltage at the terminals of the tube remains constant, but a variation in voltage takes place across the resistance.

$$\Delta E = R \Delta i$$

Now because

$$\Delta l = L \Delta i$$

$$\frac{\Delta i}{i} = \frac{\Delta l}{l} = \frac{\Delta E}{E - N}$$

The change in area of the negative glow (or its length if the electrodes are in the form of rods), is proportional to  $\Delta E$ , and can be made large by choosing  $N$  in the neighborhood of  $E$ . The glow tube oscillograph can therefore be made to amplify the voltage variations. The changes in the negative glow are readily photographed and furnish a permanent record. There is no lag up to frequencies of the order of 10,000.

### References

1. BARKHAUSEN, H. Elektronenröhren, 1929.
2. CHAUVIERRE, M. L'alimentation sur le secteur. 1928.
3. FLEMING, J. A. Propagation of electric currents, 4th ed. 1927.
4. GUNTHER-SCHULZE, A. Elektrische Gleichrichter, 2nd ed. 1929.
5. HATHAWAY, K. A. A.C. Radio Guide. 1929.
6. JEZEWSKI, M. Z. Physik, 48:123. 1928.
7. JOLLEY, L. B. W. Alternating current rectification, 3rd ed. 1928.
8. MORECROFT, J. H. Principles of radio communication, 2nd ed. 1927.
9. MUELLER, K. E. Quecksilberdampfgleichrichter. 1925.
10. PALMER, L. S. Wireless principles and practice. 1928.
11. PETERS, L. J. Theory of vacuum tube circuits. 1927.
12. Smithsonian Miscellaneous Collections, I, 74. 1922.
13. STINCHFIELD, J. M. Radio Broadcast, Aug. 1929; Proc. Inst. Radio Eng. 17: 1929.
14. Terman, F. E. Radio Broadcast, March, April, May. 1929.
15. WISE, R. Radio Broadcast, April, 1929.

# STUDIES ON REACTIONS RELATING TO CARBOHYDRATES AND POLYSACCHARIDES. XXVII. SYNTHESIS AND STRUCTURE OF TRICHLOROETHYLIDENE GLYCEROL<sup>1</sup>

BY HAROLD HIBBERT<sup>2</sup>, J. G. MORAZAIN<sup>3</sup> AND A. PAQUET<sup>3</sup>

## Abstract

A discussion is given as to the nature of the factors involved in the mechanism of the formation of cyclic acetals and ketals from carbonyl and polyhydroxy derivatives. It is shown that in the two reactions involved, namely, the primary formation of a half-acetal, followed by removal of water and ring-closure, the polarity of the carbonyl derivative plays a very important rôle. Chloral combines with polyhydroxy compounds to give exceptionally stable half-acetals, which can be converted into the cyclic derivatives by the use of strong dehydrating agents.

In contrast with the behavior of other aldehydes it is found that the condensation of chloral with glycerol yields only one ring compound, namely the five-membered derivative. No satisfactory explanation of this can, as yet, be given.

The identity of the trichloroethylidene glycerol obtained by the condensation of glycerol and chloral under the influence of concentrated sulphuric acid has been established as a five-membered cyclic acetal by a comparison of the properties of its methyl ether with those of trichloroethylidene glycerol  $\alpha$ -methyl ether prepared from glycerol  $\alpha$ -methyl ether.

## Introduction

The work described in the present paper is a continuation of the various researches on cyclic acetal formation carried out during the past decade by Hibbert and co-workers (10)\* and involves the condensation of an aldehyde or ketone with polyhydroxy compounds, especially glycols and glycerols.

The condensation of an aldehyde with a polyhydroxy derivative may be brought about, as indicated in the last communication (16), either (a) under the influence of heat and pressure and in the absence of a catalyst, (b) under the catalytic influence of an acid, or (c) by the use of a solid dehydrating agent acting in a neutral medium.

### *Mechanism of Simple Acetal Formation*

The physico-chemical work of Adkins and co-workers (1) provides definite and reliable evidence that the formation of a simple acetal from an aldehyde and an alcohol takes place in two stages:

\*The recent interesting communications by Dworzak and Lasch (8) and by Van Roon (21) in this field provide welcome confirmation of the theories developed in the earlier work of Hibbert and co-workers.

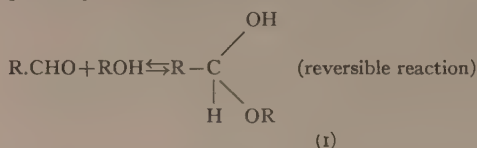
<sup>1</sup> Manuscript received January 10, 1930.

Contribution from the Industrial and Cellulose Chemistry Laboratories, McGill University, Montreal, Canada. Constructed from the theses presented by J. G. Morazain and A. Paquet in partial fulfilment of the requirements for the degree of Master of Science.

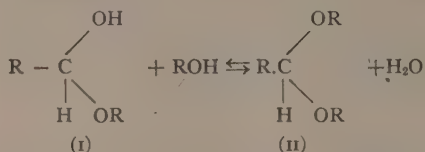
<sup>2</sup> Professor of Industrial and Cellulose Chemistry, McGill University.

<sup>3</sup> Post-graduate student, McGill University and holder of a Scholarship under the National Research Council of Canada.

## 1. The primary formation of a half-acetal:



2. The interaction of this with a second molecule of alcohol (likewise a reversible reaction):



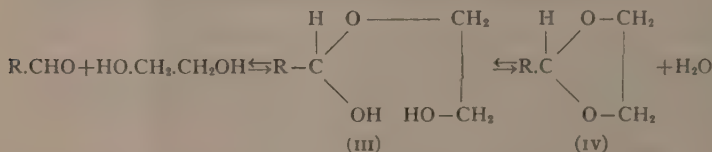
The removal of water from the sphere of action by the use of such products as anhydrous zinc chloride, calcium chloride, etc., frequently facilitates the reaction.

*Mechanism of Cyclic Acetal Formation*

Up to the present no detailed physico-chemical study of cyclic acetal formation has been made\* but the results obtained in these laboratories would seem to indicate in a broad general manner the nature of the mechanism involved.

There are two main changes taking place in all cases:

## 1. A primary reaction involving a half-acetal formation (III):



## 2. Conversion of this into a cyclic acetal (IV) by loss of water.

As in "simple" acetal formation we are concerned with chemical equilibria in both cases.

Evidence in support of these statements is to be found in the present communication and in previously published investigations.

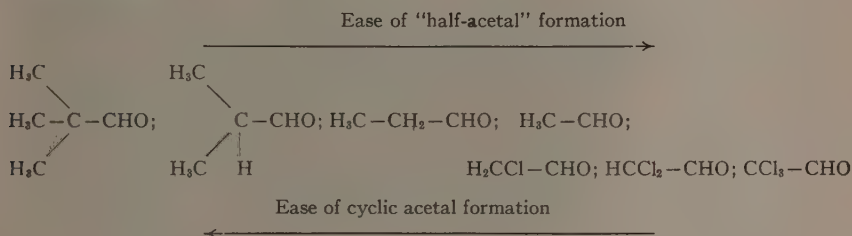
That the "primary" stage in cyclic acetal synthesis is the formation of a half-acetal is shown by the fact that in certain cases, such as the interaction between chloral and glycols, these intermediate products can be isolated, identified and readily converted into the corresponding cyclic acetal.

\*Work of this character is in progress. The previous inability to carry out such work was due to the difficulty experienced in finding a suitable type of aldehyde capable of undergoing cyclic acetal formation in the absence of a catalyst, or dehydrating agent. This difficulty has now been overcome and the mechanism of cyclic acetal formation with polyhydroxy derivative is being investigated along physico-chemical lines involving temperature, solvent, spatial relations and other physical factors.



The ease of formation of the latter will depend on a number of factors, involving (a) relative stability of the half-acetal, and (b) spatial considerations relating to polar character and ease of ring formation.

In the case of a series of aldehydes such as the alkyl- and halogen-substituted acetaldehydes:



it is to be expected, in accordance with the general principles enunciated by Arthur Michael (20) that an *increase* in the tendency towards addition of water, alcohols and polyhydroxy compounds will be found as the polar character becomes more negative, resulting in an increase in the relative stability of the half-acetal formed.

On the other hand the reverse should hold true with respect to the relative ease of ring formation. Thus the tendency towards "cyclic acetal synthesis" with respect to the same glycol should decrease from trimethyl acetaldehyde to chloral, and recent work indicates this to be the case. For example, the half-acetal from chloral and ethylene glycol is so stable that it can be recrystallized unchanged from organic solvents, and requires the use of concentrated sulphuric acid at a moderately elevated temperature to convert it into a cyclic acetal.

On the other hand, cyclic acetal formation from trimethylacetaldehyde and glycols, or glycerols, readily takes place at a relatively low temperature and in the absence of a catalyst.

### *Condensation of Aldehydes with Glycerol*

The mechanism of the condensation of aldehydes with glycerol is more complicated, inasmuch as in all cases there exists the possibility of the simultaneous formation of a five- and a six-membered cyclic acetal. It was this particular aspect of the subject, and the application of such reactions to the determination of the structure of carbohydrates and polysaccharides, which provided the original incentive to a more thorough investigation of these reactions.

It seems necessary to emphasize this in view of certain recent statements of Böseken (5) and van Roon (22). Our investigations dealing with the preparation of cyclic acetals and methods for the synthesis of the isomeric glycerol cyclic acetals, have been in progress during the last seven or eight years, and in one of the earlier papers (11) it was pointed out that "The condensation of

glycerol (1 mol) with an aldehyde (1 mol) represents an "intramolecular partition", and it seems logical to assume that in these reactions two isomers (namely, a five- and a six-membered cyclic acetal) will be formed in each case, the relative amounts being dependent on the molecular configuration of the aldehyde employed. An investigation of such reactions is being carried out."

This view, expressed by the senior author at the time, has since received abundant experimental verification at the hands of Hibbert and co-workers.

Since the first published reference by Böeseken and co-workers to such specific phenomena is to be found in the recent paper published by van Roon as well as in the latter's thesis (21) it is considered unnecessary to discuss the question of priority.

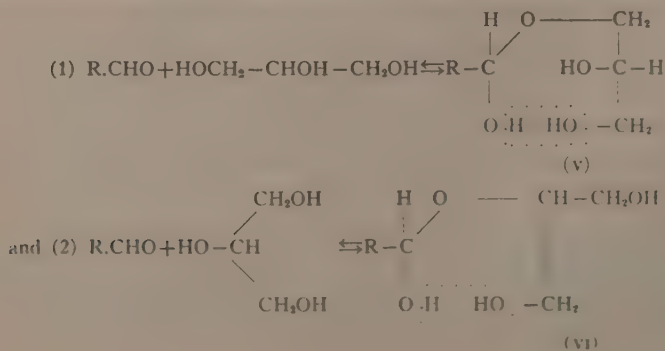
The essentially new and valuable contribution in van Roon's work is the emphasis placed on the important rôle played by temperature on the ratio of isomeric glycerol acetals formed; the importance of this factor was overlooked by Hibbert and co-workers.

The tacit assumption by earlier workers (9, 19) that, when benzaldehyde condenses with a polyhydroxy derivative such as a carbohydrate, *only* a five-membered cyclic acetal is formed, appeared directly contrary to the Michael partition principle (20). According to the latter it is to be expected that condensation of an aldehyde with a 1:2:3 trihydroxy derivative will lead to the formation of a mixture of a five- and a six-membered cyclic acetal, provided that no abnormal spatial relations exist.

In place of the more complex carbohydrates it seemed advisable to employ a much simpler derivative, namely glycerol, and the various researches carried out primarily with this object in view, namely the proof of the simultaneous formation of the five- and six-membered cyclic acetals, have yielded striking corroboration of the theory.

It is, in fact, largely as a result of this work that the use of aldehydes as well as other carbonyl condensation products for determining the structure of carbohydrates has now, to a great extent, been abandoned.

In the condensation of an aldehyde with glycerol, it is evident that the first stage of the reaction, namely, half-acetal formation, may take place in one, or both, of two ways, involving the  $\alpha$  and  $\beta$  hydroxyl groups respectively:



In the second phase, namely ring formation, it is apparent that (vi) can only give rise to a five-membered acetal, while (v) on the other hand, may form either a five- or six-membered ring, or a mixture of both.

The ease of formation of the half-acetal should follow, with respect to the aldehydes, in the same order as in the case of a 1:2 and a 1:3 glycol; this appears to be the case, and chloral forms by far the most stable intermediate.

The proportion in which the five- and six-membered rings will be formed from the half-acetal would seem to depend essentially on two factors:

- (a) The ratio of half-acetal formation from the  $\alpha$  and  $\beta$  hydroxyl groups respectively, and
- (b) Tendency and ease of ring formation as influenced by spatial, or other considerations, such as polar character of the atoms or radicals present.

Depending on the polar character of the carbonyl group, half-acetal formation involving the  $\alpha$  and  $\beta$ -hydroxyl groups may proceed at such widely varying rates that the reaction may be practically confined to only one of them, much in the same way that the action of sodium and HCl respectively on glycerol yields only  $\alpha$  sodium glycerate, and glycerol  $\alpha$  chlorhydrin, with none, or at least no appreciable quantity, of the  $\beta$  derivative.

If the assumption could be made that half-acetal formation in the case of chloral involved exclusively the  $\beta$  hydroxyl group, it is evident this would account for the experimental result found, namely exclusive formation of the five-membered acetal.

Such a conclusion, however, in the light of present available information, appears to be unwarranted.

On the other hand, if the primary addition of the chloral involves the  $\alpha$  hydroxyl group it is difficult to account for the absence of the six-membered acetal in the condensation product resulting from the action of strong sulphuric acid in the second phase of the reaction.

Apparently in this case spatial relations, and the strongly negative polar character of the  $\text{CCl}_3$ -radical, play an important rôle, but the exact nature of these cannot be settled until much further work has been carried out. That chloral can readily form six-membered acetals is shown in the ease with which cyclic acetal formation takes place with 1:3 dihydroxy derivatives such as trimethylene glycol (3) and pentaerythritol (6).

The same difference in the ease of cyclic acetal formation found between trimethyl acetaldehyde on the one hand, and chloral on the other, with respect to cyclic acetal formation from 1:2 glycols is also found in the case of glycerol. With the former aldehyde, condensation takes place readily at a moderate temperature and in the absence of a catalyst, while with the latter, condensation only takes place under the influence of a strong dehydrating agent such as sulphuric acid, zinc chloride, etc.\*

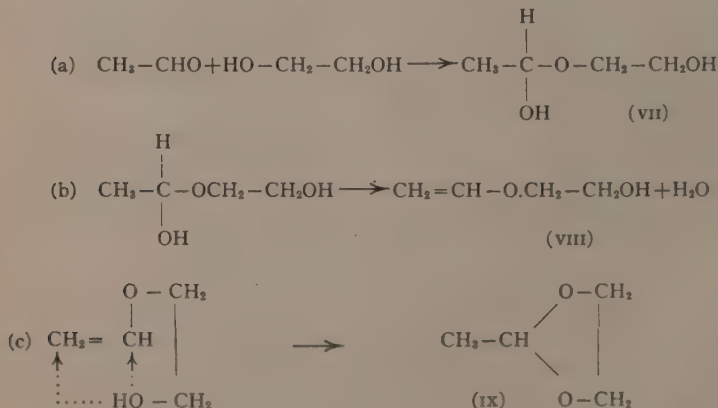
\*This work is being carried out by Mr. A. Paquet. The cyclic acetal from trimethylacetaldehyde and ethylene glycol boils at 129-130°, that from glycerol at 80.5-88°/7 mm.

The very recent work of Böeseken and van Roon (21) forms a welcome confirmation of the theories put forward in earlier investigations (13), and serves to establish the validity of Michael's partition principle as applied to the formation of glycerol cyclic acetals. Whether the views put forward by Böeseken (4) relative to the rôle exerted by the assumed spatial relations of the hydroxyl groups in the case of glycol and glycerol cyclic acetal formation have any justification is uncertain; this must await further experimental work dealing with the physico-chemical aspects of cyclic acetal- and ketal-synthesis, especially those at present under investigation, which do not involve the use of a catalyst.

The criticism by Böeseken (5) of the partition experiments carried out by Hibbert and Hill (12) with a mixture of one mol of each of two glycols and one mol of an aldehyde, in presence of dilute acid as catalyst, appears to be unwarranted. This author assumes that the experimental results in question are vitiated by the fact that no consideration was given to the possibility of a difference in the reaction velocity of the primary half-acetal formation. Such an argument would be valid only in the case of glycols having a marked difference in polar character which could scarcely be claimed of the compounds in question.

#### *Other theories of cyclic acetal formation*

The view previously expressed by Hill and Pidgeon (17) and by Hill (18) that cyclic acetal formation involves the removal of water from the "primary half-acetal", with formation of a vinyl derivative, the latter then undergoing ring formation due to the migration of a hydrogen atom:



is obviously not in accordance with the experimental facts. Aldehydes such as benzaldehyde, chloral and trimethyl acetaldehyde, which do not contain a hydrogen atom attached to the carbon atom adjacent to the carbonyl group, readily form cyclic acetals with glycols and glycerols and in such cases the proposed mechanism is evidently quite inapplicable.



### Experimental Details

#### *Hydroxy-trichloroethylidene glycol* ( $\text{CCl}_3-\text{CHOH}-\text{O}-\text{CH}_2-\text{CH}_2\text{OH}$ )

De Forcrand (7) found that on adding chloral (1 mol) to anhydrous ethylene glycol (1 mol) considerable heat was evolved, and on standing at room temperature a white crystalline product (m.p.  $42^\circ\text{C}$ .) separated out, which proved to be the glycol half-acetal,  $\text{CCl}_3-\text{CHOH}-\text{O}-\text{CH}_2-\text{CH}_2\text{OH}$ . This on treatment with concentrated sulphuric acid is readily converted into the cyclic acetal.

31 gm. of pure ethylene glycol was mixed with 83 gm. of pure anhydrous chloral, whereby considerable heat was evolved. The viscous liquid after standing for two weeks at  $0^\circ\text{C}$ . solidified completely to a white crystalline solid. This was washed several times with anhydrous ligroin, dried on a porous plate, and recrystallized several times from a mixture of anhydrous ether and anhydrous ligroin, yielding colorless crystals, m.p.  $50-51^\circ$ , that is, some  $8^\circ$  higher than found by de Forcrand. (Cl, 50.70%; calcd for  $\text{C}_4\text{H}_5\text{O}_7\text{Cl}_3$ , 50.79%).

On treatment with concentrated sulphuric acid, the half-acetal is partially converted at room temperature into the cyclic acetal, more completely however on heating at  $65-70^\circ\text{C}$ . for four hours. The cyclic acetal thus obtained melted at  $40^\circ$ . It corresponded in all of its properties with the product obtained by condensing chloral hydrate with glycol, by means of concentrated sulphuric acid. The compound was also prepared from chloral hydrate by the use of concentrated sulphuric acid as suggested by Böeseken (2). *Method*: 31 gm. of ethylene glycol (one mol) was mixed with 83 gm. of chloral hydrate (one mol) and 63 gm. of concentrated sulphuric acid added. The mixture was heated on the water bath, with stirring, for four hours at  $65-70^\circ$ . The product separated into two layers on cooling. These were separated, the upper one washed several times with water to remove the acid and glycol, and then dissolved in ether. The ethereal solution was dried over fused potassium carbonate, the ether removed by distillation, and the residual oil purified by fractionation under reduced pressure. The main product was a liquid, b.p.  $99^\circ/25\text{ mm}$ . which solidified immediately in the condenser. It was insoluble in water and after several crystallizations from a mixture of alcohol and water, melted at  $40^\circ\text{C}$ . Yield 26%. The product has a pleasant, camphor-like odor and is somewhat volatile. [Cl, 55.85, 55.76%; calcd for  $\text{C}_4\text{H}_5\text{O}_2\text{Cl}_3$ : 55.61. Mol. wt. (ethylene bromide): Found 189.5, 180; calcd 191.5].

#### *Behavior of glycerol with chloral in the absence of a catalyst*

23 gm. (one mol) of glycerol and 36 gm. (one mol) of chloral were mixed together at room temperature. Apparently a reaction occurred as the temperature rose immediately to  $80-90^\circ$ . The product was allowed to stand for two hours at room temperature. The resulting, very viscous liquid, on distillation under reduced pressure, yielded only chloral, chloral hydrate and glycerol with practically no residue. This behavior would appear to indicate that the

product is probably a half-acetal addition product. When poured into water, it gave a small amount of a white precipitate (possibly a polymer of chloral) soluble in ether, ethyl alcohol, ethyl acetate and acetone. No crystalline substance could be isolated from the crude reaction product.

In a second experiment, 23 gm. (one mol) of glycerol and 72 gm. (two mols) of chloral were allowed to react together in the hope that half of the chloral might act as a dehydrating agent, thus assisting the cyclic acetal formation, but this proved not to be the case.

*Reaction of glycerol with chloral and chloral hydrate in the presence of a mild catalyst*

23 gm. (one mol) of glycerol and 37 gm. (one mol) of chloral were mixed with one drop of concentrated hydrochloric acid, and allowed to stand at room temperature for two hours. Distillation, after neutralization, yielded the constituents unchanged.

In a second experiment, chloral hydrate was substituted for chloral. 23 gm. (one mol) of glycerol was mixed with 41 gm. (one mol) of chloral hydrate in the presence of one cc. of concentrated hydrochloric acid. The mixture was heated at 80° C. for five hours, neutralized and distilled. There was no indication of any cyclic acetal formation.

*Preparation of trichloroethylidene glycerol*

Trichloroethylidene glycerol has been synthesized by Böeseken (2) and by Yoder (23). Both authors assume, without any experimental evidence, that the product is a five-membered cyclic acetal. The following experimental data show this to be the case and at the same time indicate the absence of the six-membered cyclic acetal in the reaction product.

*Yoder's Method*

30 gm. (one mol) of glycerol and 50 gm. of chloral hydrate (one mol) were mixed, heated to 80° C., and then 25 gm. of anhydrous zinc chloride gradually added during the course of 15 to 20 minutes. After complete solution, the mixture was heated in a sealed tube at 130° for eight hours. A brown syrup was left which did not smell of chloral. It was washed with water, dissolved in ether and the ethereal solution dried over calcium chloride. On distillation, a colorless liquid, b.p. 151–5°/40 mm. was obtained, apparently identical with Yoder's product. Yield 20%.

*Böeseken's Method*

As recommended by Böeseken, 125 gm. of 95% sulphuric acid was mixed with 46 gm. (one mol) of glycerol and 90 gm. (1.1 mols) of chloral hydrate. The mixture was slowly warmed to 65–70° C. in the course of 20 minutes, and maintained there for 1.5 hours, during which a homogeneous solution had formed. The resulting liquid was cooled, poured into water and the oil allowed to settle. The latter was taken up with ether, washed with water, dried and distilled under reduced pressure. Yield: 25 gm. (20%); b.p. 126–133°/10 mm.; re-distilled, b.p. 132–133°/10 mm.

The method was later modified in the following manner: On completion of the reaction at 65–70°, the resulting mixture was cooled, extracted with chloroform, neutralized with calcium carbonate, dried over anhydrous sodium sulphate, filtered, and the solvent then distilled off. Yield 20%. This modification permits of much easier handling of the reaction mixture.

Substitution of chloral for the hydrate, under the same conditions gave a yield of 25%, the reaction proceeding somewhat more rapidly in this case.

Replacement of the concentrated sulphuric acid as catalyst by 85% phosphoric acid gave a negative result.

*Proof of the structure of the trichloroethylidene glycerol*

As indicated below, the reaction product consists entirely of 1:2-trichloroethylidene glycerol. This was proved by converting it into its methyl ether and establishing the identity of the latter with the 1:2-trichloroethylidene glycerol  $\alpha$ -methyl ether synthesized by condensing glycerol  $\alpha$ -methyl ether with chloral.

*(a) Methylation of trichloroethylidene glycerol*

25 gm. of dimethyl sulphate and a solution of 15 gm. of NaOH in 21 cc. of water were allowed to drop from separate tap-funnels into the well-stirred acetal during a period of two hours, the temperature being kept at 35–40° C. A solution of three grams of sodium hydroxide dissolved in 25 cc. of water was then added and the reaction mixture heated at 75°C. for one hour to destroy the excess dimethyl- and methyl-sodium-sulphates. The reaction product was extracted with chloroform, dried over calcium chloride and distilled. It boiled constantly at 108°/10 mm.  $n_D^{20}$ : 1.4806; mobile liquid, colorless, pleasant odor. Yield 20.6 gm. (80%).

*(b) Synthesis of trichloroethylidene glycerol  $\alpha$ -methyl ether.*

To a mixture of 15 gm. (one mol) of glycerol  $\alpha$ -methyl ether (b.p. 110°/13 mm.) (prepared from sodium methoxide and  $\alpha$ -monochlorohydrin), and 23.5 gm. (one mol) of powdered chloral hydrate, was added 18 gm. of concentrated sulphuric acid. The condensation goes more readily than with glycerol and chloral hydrate, but is accompanied by a more pronounced darkening of the product. The mixture was heated for one hour at 65°C., the reaction product extracted with chloroform, neutralized with calcium carbonate, dried over sodium sulphate, filtered and distilled. B.p. 112°/13 mm. Yield, 5 gm. (15%).  $n_D^{20}$ : 1.4803.

*Further identification of the two trichloroethylidene glycerol  $\alpha$ -methyl ethers*

Each of the two liquid  $\alpha$ -methyl ethers synthesized as above was diluted with three parts of ethyl alcohol and cooled to –70°, when crystals of the ether separated out. The one made from trichloroethylidene glycerol acetal by methylation, melted between 15–16°, quite sharply, while the other, made by the condensation of glycerol  $\alpha$ -methyl ether and chloral, melted less sharply at 11–12°. These melting points indicate that probably mixtures of the two

geometrical isomers of the same five-membered ring compound were obtained by crystallization. The absence of *two structural isomers* is indicated by the small difference in melting points.

The product is very resistant towards acid hydrolysis, which prevented its conversion into the corresponding glycerol  $\alpha$ -methyl ether for purposes of identification.

TABLE I  
PROPERTIES OF TRICHLOROETHYLIDENE GLYCEROL

	B.P.	$n_D^{20}$	$D_4^{20}$
Trichloroethylidene glycerol			
(a) From chloral and glycerol with conc. sulphuric acid	113°/0.17 mm.	1.5034	1.5425
(b) From chloral and sodium glyceroxide with conc. sulphuric acid	110°/0.125 mm.	1.5032	1.5428
Trichloroethylidene glycerol $\alpha$ -methyl ether			
(a) From the acetal	108°/10 mm.	1.4806	1.4228
(b) From glycerol $\alpha$ -methyl ether	112°/13 mm.	1.4803	1.4222

*Further attempts to isolate an equilibrium mixture of five- and six-membered trichloroethylidene glycerols.*

Since no six-membered derivative had been obtained in the previous work, some further experiments were performed in the hope of effecting its synthesis. It is evident that if glycerol reacts with chloral to form the  $\beta$ -half-acetal (of type VI) exclusively, only the five-membered acetal can be formed. On the other hand it seemed possible that if the catalytic condensing agent, for example, sulphuric acid, were allowed to react first with glycerol, an addition to the  $\alpha$ -carbon might occur, giving the sulphate  $[\text{CH}_2\text{OSO}_3\text{H}-\text{CHOH}-\text{CH}_2\text{OH}]$ . Chloral hydrate might then react with this to yield the  $\alpha$ -half-acetal, which, by ring closure, could form both the five- and six-membered isomers. No evidence however of the formation of the latter derivative was obtained.

A further experiment was carried out in which monosodium glyceroxide ( $\text{CH}_2\text{OH}-\text{CHOH}-\text{CH}_2\text{ONa}$ ) was first prepared, this then decomposed with concentrated sulphuric acid, and mixed with one mol of chloral hydrate. Here also only the five-membered acetal was obtained.

*Attempted conversion of the five-membered into the isomeric six-membered acetal.*

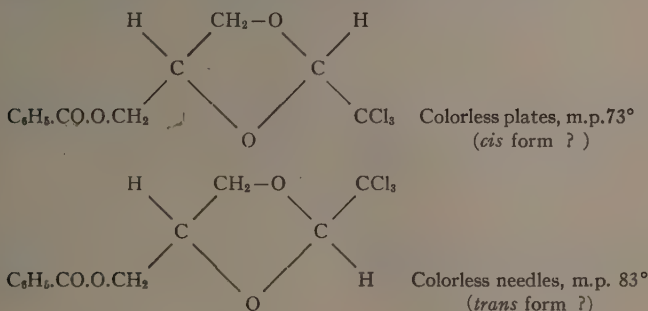
Inasmuch as the mutual interconversion of the five- and six-membered acetals under the influence of traces of acid has been shown to take place readily (15), an attempt was made to convert the five-membered trichloroethylidene glycerol into the six-membered compound under similar conditions.



Five grams of the five-membered trichloroethylidene glycerol was treated at 70° with a few bubbles of dry HCl. After standing for three hours at this temperature, with occasional shaking, the mixture was cooled and the refractive index measured. This was found to be 1.5029 at 20° as compared with 1.5032 for the pure acetal. The heating was then continued until a temperature of 120° was reached without causing any further change, except a very slight decomposition.

*Preparation of trichloroethylidene glycerol  $\alpha$ -benzoate.*

90 gm. (one mol) of the trichloroethylidene glycerol obtained as above was condensed with 54 gm. (one mol) of benzoyl chloride in the presence of 150 gm. of pyridine. After standing overnight in a warm place, the mixture was cooled and poured into one litre of cold water and stirred until the oil solidified. The yellow, amorphous mass so obtained was separated by filtration, dissolved in benzene, washed with a cold saturated solution of sodium bisulphate, then with a cold solution of sodium bicarbonate and finally dried over calcium chloride. Crystallization was effected by addition of ligroin and cooling, two different crystalline products being obtained, which, judging by the proximity of their melting-points, evidently were the *cis* and *trans* forms.



The two types of crystals were separated by hand. [Lower melting isomer, Cl, found, 32.1%; calcd for  $C_{12}H_{11}O_4Cl_3$ , 32.7%. Mol. wt. (ethylene bromide), found 306.1; calcd, 325.5. Higher melting isomer, Cl, found 32.0%; calcd, 32.7%. Mol. wt. (ethylene bromide), found 316.6; calcd, 325.5.]

A solution of the benzoate in dry toluene was inert towards sodium, indicating absence of free hydroxyl groups. These facts, it is true, do not definitely prove that each of the benzoates is a derivative of the five-membered isomer. However similar researches on other acetals (14) have shown that geometrical isomers of this type have melting points differing by only a few degrees, while in the case of structural isomers this difference is much greater. The choice of the above formulæ would thus seem to be justified.

Attempts to isolate from the crude reaction product a third and a fourth benzoate, corresponding to the six-membered ring acetal, were unsuccessful.

*Preparation of trichloroethylidene glycerol p-nitro-benzoate*

The mechanical separation of the crystalline benzoates was difficult to carry out and it was hoped to obtain better results with the *p*-nitrobenzoates. Using the same conditions, two different crystalline geometrical isomers were also obtained, which, after mechanical separation by hand, were found to melt at 102° and 102.5° respectively, a mixture of equal parts melting around 92°.

### Conclusions

1. The theory of cyclic acetal formation is discussed and certain factors indicated as having an important bearing on the mechanism of such reactions.
2. The polar character of the aldehydes exerts an important influence as shown by the ease of formation of cyclic acetals from trimethyl acetaldehyde, in the absence of a catalyst, as compared with the necessity for the use of a strong dehydrating agent such as sulphuric acid, in the case of chloral.
3. The replacement of the hydrogen atoms in the methyl radicle of acetaldehyde by methyl groups apparently decreases both the tendency towards half-acetal formation and stability of the half-acetal but increases the ease of ring formation. The reverse appears to be the case when the chlorine atoms, as in chloral, are substituted for the hydrogen atoms in the methyl radicle.
4. No satisfactory explanation can at present be offered for the fact that chloral and acetone, when condensed with glycerol, yield only the five-membered acetal and ketal respectively. Apparently physico-chemical spatial relations, involving polar character, play an important rôle but the exact nature of these awaits further investigation.
5. The structure of the cyclic acetal formed by condensing glycerol with chloral has been definitely established by the synthesis of its methyl ether and comparison of this with that of the condensation product obtained from chloral and glycerol  $\alpha$  methyl ether.
6. No evidence of the presence of a six-membered acetal could be found in the condensation product from chloral and glycerol.
7. Chloral has been shown to combine readily with ethylene glycol yielding a very stable crystalline half-acetal, which when warmed with concentrated sulphuric acid is converted into the corresponding cyclic acetal. On the other hand dilute acids decompose the half-acetal into its constituents.
8. Trichloroethylidene glycerol is exceedingly stable towards hydrolysis. It is neither attacked by water, aqueous alcoholic solution of HCl, nor by aqueous 35% HCl, but concentrated sulphuric acid gives rise to some free chloral.

## References

1. ADKINS, H. and BRODERICK, A. E. J. Am. Chem. Soc. 50:499. 1928.
2. BÖESEKEN, J. Verslag. Akad. Wetensch. Amst. 35:1084. 1926.
3. BÖESEKEN, J. Verslag. Akad. Wetensch. Amst. 35: 1089. 1926.
4. BÖESKEN, J. Bull. Soc. Chim. Belg. 37:383. 1928.
5. BÖESEKEN, J. Rec. trav. chim. 48:931. 1929.
6. BÖESEKEN, J. and FELIX, B. B. C. Ber. 61:787. 1928.
7. DE FORCRAND, R. Compt. rend. 108:618. 1889.
8. DWORZAK, R. and LASCH, T. M. Monats. 51:59. 1929.
9. HAWORTH, W. N. Chem. Soc. Ann. Reports pp. 74-92. 1926.
10. HIBBERT, H. and Co-WORKERS. J. Am. Chem. Soc. 1921-1929.
11. HIBBERT, H. and HILL, H. S. J. Am. Chem. Soc. 45:748. 1923.
12. HIBBERT, H. and HILL, H. S. J. Am. Chem. Soc. 45:3108. 1923.
13. HIBBERT, H. and Co-WORKERS. J. Am. Chem. Soc. 50:2235, 2242, 3120. 1928.
14. HIBBERT, H. and CARTER, N. M. J. Am. Chem. Soc. 50:3376. 1928.
15. HIBBERT, H., PLATT, M. and CARTER, N. M. J. Am. Chem. Soc. 51:3641. 1929.
16. HIBBERT, H. and MORAZAIN, J. G. Can. J. Research 2:35. 1930.
17. HILL, H. S. and PIDGEON, L. M. J. Am. Chem. Soc. 50:2718. 1928.
18. HILL, H. S. J. Am. Chem. Soc. 50:2725. 1928.
19. IRVINE, J. C. and SCOTT, J. P. J. Chem. Soc. 103:575. 1913.
20. MICHAEL, A. Ber. 39:2138, 2143, 2149, 2153, 2157, 2569, 2785, 2789. 1906; 40:140. 1907.  
J. prakt. Chem. 60:339, 341, 437. 1899.
21. VAN ROON, J. D. Diss. Delft, 1929. Rec. trav. chim. 48:173. 1929. (C.A. 23:1874. 1929).
22. VAN ROON, J. D. (C.A. 23:1875. 1929).
23. YODER, L. J. Am. Chem. Soc. 45:475. 1923.

# THE OXIDATION OF XANTHATES AND SOME NEW DIALKYL SULPHUR- AND DISULPHUR- DICARBOTHIONATES<sup>1</sup>

BY ADRIEN CAMBRON<sup>2</sup> AND GEORGE STAFFORD WHITBY<sup>3</sup>

## Abstract

Sodium tetrathionate and cyanogen bromide in water have a specific oxidizing action on alkali xanthates, the latter being smoothly converted to xanthogen disulphides without the formation of by-products. Cyanogen bromide in alcohol converts xanthates into xanthogen monosulphides. Other new reagents for the oxidation of xanthates to xanthogen disulphides under certain conditions are—nitrous acid, nitrosyl chloride, chloramine-T, benzene sulphonyl chloride.

The following new xanthogen disulphides, viz., *n*-hexyl (liquid),  $\beta$ -ethoxyethyl (liquid),  $\beta$ -phenylethyl (m.p. 49.5–50°), and the following new xanthogen monosulphides, viz., *n*-hexyl (liquid), benzyl (m.p., 82.5°),  $\beta$ -ethoxyethyl (liquid), are described.  $\beta$ -phenylethyl xanthogenamide (m.p., 79°) is also described.

Pyrolysis of  $\beta$ -phenylethyl xanthogen disulphides yields styrene.

## Introduction

In a previous paper (16) it was shown that the disulphides obtained by the oxidation of xanthic acids can be converted into the corresponding monosulphides by the action of alkali cyanides. In a subsequent paper in this Journal the mechanism of this conversion will be discussed. The present paper describes some new methods of oxidizing salts of xanthic acids to the corresponding disulphides, and incidentally describes new members of this series of disulphides and of the corresponding series of monosulphides.

The sulphides in question, (RO.CS)<sub>2</sub>S<sub>2</sub> and (RO.CS)<sub>2</sub>S, have been designated in the literature in various ways. Here they are designated systematically as dialkyl disulphurdicarbathionates and dialkyl sulphurdicarbathionates respectively\*, and trivially as xanthogen di- and mono-sulphides respectively.

The reagents commonly used hitherto for the oxidation of xanthates to xanthogen disulphides have been as follows: (a) Iodine in alcoholic or in aqueous potassium iodide solution, (3, 5), the reaction being:



(b) Copper sulphate (1, 2), the reaction being:



(c) Chlorine (4, 10). Chlorine is probably the least satisfactory of these agents, the products being very impure (16). And, as experimental results to be described later will serve to show, even the other two agents are not well adapted to the preparation of pure xanthogen disulphides, except in the limited number of cases where the products are solids which can be purified by recrystallization.

\* Following a recommendation of Dr. C. Smith, Editor, *Journal of the Chemical Society*.

<sup>1</sup> Manuscript received February 3, 1930.

Contribution from the Department of Chemistry, McGill University, Montreal.

<sup>2</sup> Associate Research Chemist, National Research Laboratories, Canada.

<sup>3</sup> Director, Division of Chemistry, National Research Laboratories, Canada.



New and improved oxidizing agents were found to be sodium tetrathionate and cyanogen bromide, both of which gave very pure products when applied to alkali xanthates. In addition, a number of other new oxidizing agents for this reaction were found, viz., nitrous acid, chloramine-T, nitrosyl chloride, benzene sulphonyl chloride. None of these latter proved to be as valuable as the two first-mentioned reagents for the preparation of pure xanthogen disulphides, but they are of some theoretical interest, and their action on xanthates represents new reactions.

Cyanogen bromide yields xanthogen disulphides from alkali xanthates only when reaction is carried out in a medium, such as water or benzene, in which either the disulphide or alkali cyanide is insoluble; in alcohol reaction proceeds through a further stage, and the product is the xanthogen monosulphide.

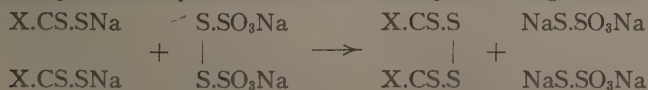
#### *Sodium tetrathionate*

By the action of sodium tetrathionate on a number of aromatic mercaptides, Footner and Smiles (6) obtained the corresponding disulphides in practically theoretical yield. It was thought that this mild oxidizing agent, neutral in reaction and water-soluble, might have a similar specific action on alkali xanthates. This was found to be the case, good yields of pure xanthogen disulphides being obtained by its use.

Benzyl xanthogen disulphide is described as an oil by Ragg (9), who prepared it by the action of copper sulphate on sodium benzyl xanthate. By the use of sodium tetrathionate, the compound was secured directly as a crystalline solid, which, after one recrystallization from glacial acetic acid, melted at 54°C. Methyl xanthogen disulphide is described as a liquid by Desains (5, p. 504), who prepared it by the action of iodine on sodium methyl xanthate, and by Ragg (9), who prepared it by the action of copper sulphate on the same salt. Whitby and Greenberg (16) secured the compound in solid form, melting at 22-22.5°C., by cooling, in carbon dioxide-ether, the liquid obtained by the use of iodine. In the present work, by the use of sodium tetrathionate, the compound was obtained directly as a solid, melting at 23-23.5°C.

It was also found that sodium tetrathionate can be used with fair results for the oxidation of salts of dithiobenzoic and dithionaphthoic acids to the corresponding disulphides.

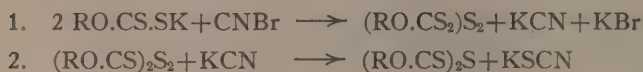
The following equation represents the mode of action of sodium tetrathionate on the compounds in question, sodium thiosulphate being formed:



#### *Cyanogen bromide*

The oxidizing action of cyanogen bromide on disubstituted dithiocarbamates was noted by von Braun and Stechele (14), who obtained tetramethyl thiuram disulphide from dimethylammonium dimethyldithiocarbamate in alcohol at a low temperature. On treating xanthates with cyanogen bromide in alcoholic

solution, the present authors found that the product was almost exclusively the xanthogen monosulphide, even when the reaction was carried out at  $-5^{\circ}\text{C}$ . The difference in the nature of the product in the two cases was found to be due to the fact that, although the alkali cyanide formed in the first stage of the reaction has no appreciable reaction on thiuram disulphides in alcohol at low temperatures, it readily desulphurizes xanthogen disulphides under the same conditions. The action of cyanogen bromide on xanthates in alcohol proceeds in two stages, as follows:



As already mentioned, the reaction represented by the second equation has been shown by Whitby and Greenberg (16) to take place in alcoholic solution.

The action of cyanogen bromide on alkali xanthates in alcoholic solution is a convenient method of preparing xanthogen monosulphides in a pure state, and does not, unlike the method of Whitby and Greenberg, involve the preparation and isolation of the disulphide as a separate step. Three new xanthogen monosulphides, viz., the *n*-hexyl, benzyl and  $\beta$ -ethoxyethyl compounds, were prepared by this method.

It was then found that reaction between cyanogen bromide and xanthates could be stopped at the first stage by using as reaction media solvents in which either the xanthogen disulphide or the alkali cyanide produced are insoluble, such as water, ether, benzene. Of these, water seemed to give the best results; in it cyanogen bromide appeared to be very specific in its action on xanthates, no indications of side reactions being observed. The preparation of  $\beta$ -phenylethyl xanthogen disulphide may be mentioned, as an example of the results with cyanogen bromide. Several samples of this compound were prepared by the iodine method, but, even under the most carefully controlled conditions, the product was invariably an oil, which failed to crystallize when cooled in carbon dioxide-acetone. It was believed, however, by analogy, that the compound, which is not described in the literature, should be a solid. And, on using cyanogen bromide for its preparation, it was secured as well-defined crystals, melting at  $49.5\text{--}50^{\circ}\text{C}$ .

The following equation, viz.,



probably correctly represents the initial and final states of the reactants, but is probably misleading as to the actual mechanism of the reaction, because the CN residue is represented as playing the same rôle as the halogen atom; this cannot be the case, since the CN residue is positive to the RO.CS.S residue (at least when R is an alkyl residue), or, in other words, since the electron affinity of the CN group is weaker than that of the RO.CS.S group and since consequently the former group cannot replace the latter from its salts\*. A more acceptable

\* Cyanogen was found experimentally to be without action on xanthates.

view as to the mechanism of the above action is to assume that cyanogen bromide first hydrolyzes in the manner shown by Nef (8), the hypobromous acid formed being the actual oxidizing agent.—



#### *Nitrous acid*

This reagent was found to be convenient for the oxidation of certain xanthates, *e.g.*, sodium ethyl and sodium isopropyl xanthates, to the corresponding xanthogen disulphides. Its use is limited, because the free xanthic acid is liberated in the course of the reaction, and hence, apparently, pure products are obtained only in those cases where the xanthic acids, such as those derived from the lower aliphatic alcohols, are sufficiently stable to undergo the oxidation process without any marked decomposition.

The reaction between xanthic acids and nitrous acid is clearly analogous to the action of the latter on hydriodic acid, as the following equations will indicate:



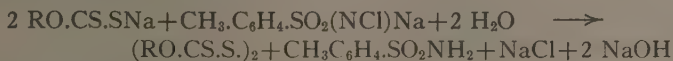
#### *Nitrosyl chloride*

The action of nitrosyl chloride on sodium isopropyl xanthate in aqueous solution gave the xanthogen disulphide. The yield (48%) was relatively low, due obviously to the readiness with which nitrosyl chloride decomposes in aqueous solution. The reaction is probably as follows:



#### *Chloramine-T*

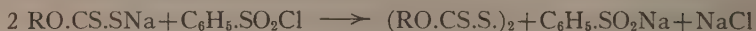
In view of the similarity in reactivity of the chlorine in chloramine-T and that in hypochlorous acid, it was expected that the former would be capable of oxidizing xanthates to xanthogen disulphides. This expectation was confirmed experimentally. But it was found, further, that, whereas the action of the hypochlorous acid formed when chlorine is passed into an aqueous solution of a xanthate, is always accompanied by side reactions, due mainly to liberation of the free acid in the course of the reaction, the action of chloramine-T appeared to be more specific, the xanthogen disulphides produced being free from by-products. The action of chloramine-T on xanthates is probably as follows, taking place in accord with the lines suggested by Schiemann and Novak (11) in a general consideration of chloramine-T as an oxidizing agent:



#### *Benzene sulphonyl chloride*

This reagent was used by von Braun and Kaiser (15) for the oxidation of dithiocarbamates to the corresponding thiuram disulphides. It was found to have a similar action on xanthates, ethyl xanthogen disulphide being obtained

in good yield by its action on sodium ethyl xanthate in ether. The following equation probably represents the action:



*Heat decomposition of  $\beta$ -phenylethyl xanthogen disulphide*

A method of dehydrating alcohols, first applied by Tschugaeff (13) to the preparation of menthene, consists in heating the methyl ester of the corresponding xanthic acid. Following such a procedure, Nametkin (7) obtained stilbene by the heat decomposition of the methyl ester of benzyl xanthic acid, the reaction being as follows:



Since dehydration in the case of phenylethyl alcohol can be intramolecular, it might be expected that styrene would be the main heat decomposition product from phenylethyl xanthates and xanthogen sulphides. And it was in fact found, on subjecting phenylethyl xanthogen disulphide to decomposition by heat that styrene was produced. The reaction can be expressed as follows:—



## Experimental

### I. ACTION OF SODIUM TETRATHIONATE ON XANTHATES

*Methyl xanthogen disulphide, (Dimethyl disulphurdicarbothionate)*  
( $\text{CH}_3\text{O.CS.S.}$ )<sub>2</sub>

30 gm. pure potassium methyl xanthate in 150 cc. water was treated with a 5% excess of sodium tetrathionate, reaction being carried out at a  $-5$  to  $-3^\circ\text{C}$ . (In all the work sodium tetrathionate was freshly prepared before each experiment). The xanthogen disulphide (98% yield) was obtained as a solid. It was recrystallized by making a saturated solution at room temperature in methyl alcohol and then cooling in an ice bath. Small prisms, m.p.  $23-23.5^\circ\text{C}$ .

*Benzyl xanthogen disulphide, (Dibenzyl disulphurdicarbothionate)*  
( $\text{C}_7\text{H}_7\text{O.CS.S.}$ )<sub>2</sub>

Sodium benzyl xanthate in water was treated with sodium tetrathionate at room temperature, the xanthogen disulphide being obtained directly as a nearly colorless solid, melting at  $46-49^\circ\text{C}$ . After three recrystallizations from glacial acetic acid, the product showed a constant melting point at  $54-54.5^\circ$ . S, 35.82% (Calcd, 35.00%).

*$\beta$ -Ethoxyethyl xanthogen disulphide, (Di- $\beta$ -ethoxyethyl disulphurdicarbothionate)*  
( $\text{C}_2\text{H}_5\text{O.CH}_2.\text{XH}_2\text{O.CS.S.}$ )<sub>2</sub>

An attempt was made to prepare the sodium derivative of the monoethyl ether of ethylene glycol by the action of metallic sodium on the alcohol, but it appeared that side reactions occurred, for a very dark-colored product was obtained on the addition of carbon disulphide. Better results were obtained by adding, with good stirring, finely pulverized sodium hydroxide (from sodium)



to a solution of excess carbon disulphide in twice the theoretical amount of the alcohol. The sodium xanthate crystallized out on standing; it was filtered off and recrystallized from acetone. On oxidizing 30 gm. of the xanthate dissolved in water with sodium tetrathionate, the xanthogen disulphide was obtained, after washing with water and drying, as a pale yellow oil.  $n_D$ , 1.5751 at 25° C. S, 38.37% (Calcd, 38.84%).

*β-Ethoxyethyl xanthogen monosulphide, (Di-β-ethoxyethyl sulphurdicarbothionate)*  
(C<sub>2</sub>H<sub>5</sub>O.CH<sub>2</sub>.CH<sub>2</sub>O.CS)<sub>2</sub>S

This was prepared from the corresponding disulphide (*supra*) by the method described by Whitby and Greenberg (16), *i.e.*, by treatment with potassium cyanide in alcohol. Reddish yellow oil;  $n_D$ , 1.5561 at 25.5° C.

#### *Dithiobenzoyl and dithionaphthoyl disulphides*

In each case the lead salt of the dithio acid was first decomposed by dilute hydrochloric acid in the presence of ether; the ether extract was washed with cold water, and the free acid converted into the sodium salt by shaking the ethereal solution with the theoretical amount of 10% sodium hydroxide solution. The solution of the sodium salt was then treated with a slight excess of a solution of sodium tetrathionate, and the precipitated disulphide was filtered off and recrystallized. The products obtained from the dithio acids did not appear to be as pure as those obtained from xanthates, as several recrystallizations were required before they showed the correct melting point.

## II. ACTION OF CYANOGEN BROMIDE ON XANTHATES

*n-Hexyl xanthogen disulphide, (Di-n-hexyl disulphurdicarbothionate)*  
(C<sub>6</sub>H<sub>13</sub>O.CS.S)<sub>2</sub>

The sodium alcoholate was first prepared by the action of metallic sodium (15gm) on excess of *n*-hexyl alcohol (100 gm) in 200 cc. toluene. The mixture was boiled under reflux, for 20 hr., until all the sodium had gone into reaction. Excess of carbon disulphide was then added with good stirring to the cooled alcoholate solution; the sodium *n*-hexyl xanthate was filtered off and was freed from excess alcohol by washing with dry ether. 40 gm. of the xanthate in 150 cc. water was treated with 10 gm. cyanogen bromide, also in water. The xanthogen disulphide separated as an oil, on the addition of the cyanogen bromide. After washing several times with water and drying, the product consisted of a pale yellow oil, with an odor faintly recalling that of the alcohol, and boiling with slight decomposition at 170-5°/3mm.  $n_D$ , 1.5569 at 25° C. S, 35.62% (Calcd, 36.21%).

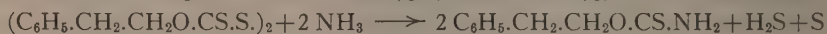
*β-Phenylethyl xanthogen disulphide, (Di-β-phenylethyl disulphurdicarbothionate)*  
(C<sub>6</sub>H<sub>5</sub>.CH<sub>2</sub>.CH<sub>2</sub>O.CS.S)<sub>2</sub>

This compound was prepared in the same manner as the *n*-hexyl compound described above, metallic sodium being used for preparing the xanthate. The xanthogen disulphide was obtained as a colorless crystalline solid, which, after several recrystallizations from glacial acetic acid, showed a constant melting point of 49.5-50° C. S, 32.52% (Calcd, 32.39%).

On distilling 25 gm. of the product under atmospheric pressure and fractionating the distillate twice, 3 gm. of a fraction boiling at 142-8° C. was collected. On treating this fraction in glacial acetic acid with slightly less than the amount of bromine required to produce a color, 1.7 gm. of a crystalline precipitate was obtained and was identified as styrene dibromide by its melting point of 73-4° C. after recrystallization from glacial acetic acid.

*$\beta$ -Phenylethyl xanthogenamide*,  $C_6H_5.CH_2.CH_2O.CS.NH_2$

This compound, which is not given in the literature, was prepared by the action of alcoholic ammonia on the corresponding xanthogen disulphide, which is described above, following the method of Debus (3) for the preparation of ethyl xanthogenamide. It was obtained in the form of large, lustrous plates from alcohol. M.p. 79° C. S, 18.49% (Calcd, 17.62%). The reaction is,

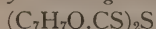


*Isopropyl xanthogen disulphide*, (*Diisopropyl disulphurdicarbothionate*)



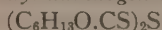
A solution of 5 gm. cyanogen bromide in 25 cc. water was added, without cooling, to a solution of 20 gm. potassium isopropyl xanthate in 75 cc. water. The product came down, in practically quantitative yield, as a pale yellow solid, which, after one recrystallization from alcohol, melted at 57.5-58.5° C. Whitby and Greenberg (16) give the melting point of the pure compound as 58.5° C.

*Benzyl xanthogen monosulphide*, (*Dibenzyl sulphurdicarbothionate*)



5 gm. cyanogen bromide and 23 gm. sodium benzyl xanthate were dissolved in 80 cc. of 80% alcohol. The solution was stirred for one hour in an ice bath, and the product was precipitated by the addition of water and extracted with ether. After evaporating off the ether, and recrystallizing the product from glacial acetic acid, the latter was obtained in the form of fine, pale yellow crystals, melting at 82.5-83° C. S, 28.00% (Calcd, 28.80%).

*n-Hexyl xanthogen monosulphide*, (*Di-n-hexyl sulphurdicarbothionate*).



Five grams cyanogen bromide and 18 gm sodium *n*-hexyl xanthate in 75 cc. alcohol were allowed to stand at room temperature for 1.5 hr. The monosulphide was then precipitated with water. After washing and drying, it was obtained in almost quantitative yield as a deep yellow oil.  $n_D^{25}$ , 1.5410 at 25° C. S, 30.48% (Calcd, 29.80%).

*Ethyl xanthogen monosulphide*, (*Diethyl sulphurdicarbothionate*)  $(C_2H_5O.CS)_2S$

A solution of 10.5 gm. cyanogen bromide in 30 cc. of 95% alcohol was gradually added to a solution of 30 gm. potassium ethyl xanthate in 100 cc. alcohol at about 0°. The product was then precipitated by the addition of water; was filtered off, and, after recrystallization from alcohol, was obtained as a yellow crystalline powder, melting at 52.5° C. (The corresponding disulphide melts at 32° C.)

## III. ACTION OF NITROUS ACID ON XANTHATES AND DITHIOCARBAMATES

*Ethyl xanthogen disulphide (Diethyl disulphurdicarbothionate) (C<sub>2</sub>H<sub>5</sub>O.CS.S.)*

A solution of 16 gm. potassium ethyl xanthate and 7.2 gm. sodium nitrite in 75 cc. water was treated gradually at 5° C. with 40 cc. of 25% sulphuric acid, with mechanical stirring. The xanthogen disulphide separated as a colorless, crystalline solid, which, after one recrystallization from alcohol, melted, correctly, at 32.5° C. The yield was 10.5 gm., or 87% of theory based on xanthate used. On repeating the experiment with hydrochloric instead of sulphuric acid, the yield of ethyl xanthogen disulphide was 97%, 14.6 gm. being obtained from 20 gm. potassium ethyl xanthate and 10 gm. sodium nitrite.

*Isopropyl xanthogen disulphide, Diisopropyl disulphurdicarbothionate) (C<sub>3</sub>H<sub>7</sub>O.CS.S.)<sub>2</sub>*

Following the procedure just described, 50 gm. sodium isopropyl xanthate and 20 gm. sodium nitrite gave 30.1 gm. (80%) of the disulphide, when sulphuric acid was used, and 32.5 gm. (87%), when hydrochloric acid was used.

*n-Butyl xanthogen disulphide, (Di-n-butyl disulphurdicarbothionate) (C<sub>4</sub>H<sub>9</sub>O.CS.S.)<sub>2</sub>*

Using sodium *n*-butyl xanthate, sodium nitrite and sulphuric acid, and following the procedure described above, for the ethyl compound, this compound was obtained as a pale yellow oil, in 80% yield.

Qualitative experiments were carried out on the treatment of sodium isoamyl and benzyl xanthates with nitrous acid, but the products were contaminated by appreciable amounts of the corresponding alcohols, as was shown by extraction of the crude xanthogen disulphides obtained with methyl alcohol.

*Tetraethyl thiuram disulphide*

Following the procedure described for the preparation of ethyl xanthogen disulphide by means of nitrous acid, 45 gm. sodium diethyldithiocarbamate and 18 gm. sodium nitrite in 100 cc. water were treated at 0° C. with 80 cc. of 25% hydrochloric acid. The thiuram disulphide separated in an almost pure state (m.p., 68.5° C.). After one recrystallization from alcohol, it showed the correct melting point, viz., 71° C.

## IV. ACTION OF NITROSYL CHLORIDE ON XANTHATES

To a chloroform solution of nitrosyl chloride, prepared according to the method of Stanek (12), was added, with good stirring and cooling, less than the theoretical amount of sodium isopropyl xanthate in aqueous solution. The chloroform layer was then separated and the chloroform removed under reduced pressure. The residue was found to consist of pure isopropyl xanthogen disulphide, melting at 57.5° C. From 5 gm. sodium isopropyl xanthate was obtained 2.1 gm. (48%) of the disulphide.

## V. CHLORAMINE-T

By adding a solution of 4 gm. sodium isopropyl xanthate to an aqueous solution containing an excess of the chloramine, isopropyl xanthogen disulphide was obtained in practically quantitative yield. After one recrystallization from alcohol, the product melted at 58.5° C.

## VI. BENZENE SULPHONYL CHLORIDE

32 gm. potassium ethyl xanthate was gradually added to a solution of 8.6 gm. benzene sulphochloride (0.5 mol.) in 100 cc. of dry ether, the reaction being carried out at -10° C. Potassium chloride was removed by filtration, the ether was evaporated off under the pump, and the pale yellow oil which remained was dissolved in alcohol. On cooling the alcoholic solution in a freezing mixture, 11.5gm. of solid ethyl xanthogen, melting at 32-32.5° C., was obtained. That the product was in fact ethyl xanthogen disulphide was confirmed by converting a portion of it into the corresponding xanthogen monosulphide by the action of potassium cyanide in alcohol.

## References

1. BILLMANN, E. Ber. 35:2185. 1902.
2. COUERBE, J. P. Ann. Chim. (2), 61:247. 1837.
3. DEBUS, H. Ann. 72:4. 1849.
4. DEBUS, H. Ann. 82:261. 1852.
5. DESAINS, P. Ann. Chim. (3), 20:498. 1847; Jahresb. 48:690. 1847.
6. FOOTNER, H. B. and SMILES, S. J. Chem. Soc. 127:2888. 1925.
7. NAMETKIN, S. and KURSANOFF, D. J. prakt. Chem. 112:165. 1926.
8. NEF, J. U. Ann. 287:315. 1895.
9. RAGG, M. Chem-Ztg. 34:92. 1910.
10. ROTHMANN, A. and STEIN, H. U.S. Patent 1,626,558. 1926.
11. SCHIEMANN, G. and NOVAK, P. Z. angew. Chem. 40:1032. 1927.
12. STANEK, V. Z. physiol. Chem. 46:263. 1905.
13. TSCHUGAEFF, L. Ber. 32:3332. 1904.
14. VON BRAUN, J. and STECHELE, F. Ber. 36:2275. 1903.
15. VON BRAUN, J. and KAISER, W. Ber. 56:549. 1923.
16. WHITBY, G. S. and GREENBERG, H. Trans. Roy. Soc. Can. 23:21. 1929.



# THE TRANSFUSION OF MATTER FROM ONE SOLID TO ANOTHER UNDER THE INFLUENCE OF HEAT—

## A NEW FACTOR IN THE PROCESS OF METAMORPHISM<sup>1</sup>

BY FRANK D. ADAMS<sup>2</sup>

### Abstract

This paper describes a case of the transfusion of material from one solid to another under the influence of heat, which, it is suggested, may throw light upon certain metamorphic processes at work in the earth's crust.

### Introduction

While the widespread action of metamorphism is coming to be recognized as playing a more and more important rôle in connexion with the processes which have been at work in giving to the older and deeper seated rocks of the earth's crust the characters which they now present, the exact nature of the forces which have been active in producing these changes is still very imperfectly understood. Much has been learned concerning the action of heat in bringing about diagenesis and most metasomatic alterations have been attributed to the action of waters or of vapors given off by cooling igneous magmas, while the studies of Sederholm and others have thrown much light on the development of "migmatites" or mixed rocks formed by the intimate penetration and partial absorption of rocks by intrusive magmas.

That however the underlying causes of metamorphism in its enormous mass action, as seen in the Archean of many regions, are still obscure and that there are probably important factors at work which as yet remain unrecognized, must be evident to all who have attempted to unravel the structures of such regions and to account for the development of the rock types which they now present. This is illustrated in a recent paper by T. T. Quirke (5) setting forth some results of an extended study which he has made of a portion of the Canadian Shield north of Lake Huron. He writes as follows:—

"In the region to the east of Killarney all the Huronian formations have been recognized in whole or in part, although altered profoundly by metasomatic regional metamorphism. Every sediment has naturally undergone a different course of changes, each of which is being treated in another work now in preparation; but this paper treats only of the alteration of the Lorrain quartzite. Nevertheless, it may be noted here that the end product of the alteration of all the sediments of the Huronian series, so far studied, seems to approach a common type, a granitic gneiss low in dark minerals.

"The means by which such changes may have been brought about are not known. No satisfactory physical-chemical explanations seem available. There appear to be no pegmatitic feeders penetrating the gneisses near

<sup>1</sup> Manuscript received March 26, 1929.

<sup>2</sup> Emeritus Vice-Principal of McGill University, Montreal, Canada, and member of the National Research Council of Canada.

Killarney, as the only veins present are numerous, very thin veins of quartz, free of feldspar. There is no indication of an interleaving of granitic bands with less granitic bands which might represent intrusions and the original country rock. The transfer of material appears to have been accomplished without the usual mechanical details."

It would be of interest to know the nature of the unknown causes which bring about results such as these—results which field studies in other pre-Cambrian areas would seem to indicate are not confined to the area studied by Dr. Quirke.

Professor T. C. Chamberlin (1) in order to account for the deformation or "flow" of solid rock which takes place under great pressure in the deeper parts of the earth's crust has supposed that where there has been no breaking or "flow of fragments", the movement may have taken place by what he terms "idiomolecular" movement:—

"Solid flow, in the sense of flow by reorganization, is really a form of growth following disaggregation. It is a serial reorganization, one integer after another changing its place and attachments. The action is individualistic (idiomolecular) in distinction from communistic (all-together) action, such as takes place when assemblages of molecules move pell-mell (liquid-flow), or by heterogeneous tumbling, as in the flow of sand and talus blocks (fragmental flow).

"In the variety of solid flow which has been called idiomolecular, the main mass of the body remains stationary and solid while individual atoms or molecules that happen to be under special stress because of their positions at angles or otherwise are forced from their points of attachment and given sufficient projection to penetrate the surrounding field of force far enough to find new attachments and again become a part in the structure. By this shift, the most intense stresses are relieved, and the new attachments help to adapt the structure to the remaining stress. Far from really being a paradoxical action, it is a specially orderly, soldier-like action to relieve the outpost under greatest pressure."

It would be a matter of great interest to know whether atoms or molecules under the conditions in question are really torn from their bases and shot through the rock to new positions of attachment, and whether movements in a solid rock can thus be accounted for. The explanation however is still in the region of hypothesis and is in any event confined to movements of rock under great pressure.

### Experimental

The writer's attention has recently been drawn to an experiment which seems to be suggestive of certain processes which may play a part—perhaps an important part—in metamorphic changes and which have not hitherto been recognized by geologists.

The Magnesite Committee of the National Research Council of Canada has for some time past been carrying out an investigation on the utilization of Canadian magnesite for certain industrial purposes and the experiment to

which reference has been made was one of a series to determine the effect of certain substances when added to magnesite clinker (dead-burned magnesite) upon the properties of magnesite bricks produced from the mixture by moulding and subsequently burning at a high temperature.

Two bricks were burned side by side in a laboratory kiln. One brick (No. 5, Fig. 1) was composed of dead-burned Austrian magnesite, low in lime, and the other (No. 6) of a synthetic clinker made from dolomitic magnesite and mill scale (largely  $\text{Fe}_3\text{O}_4$ ). The analyses of these bricks are given in Table I.

The temperature was raised by approximate increments of  $30^\circ \text{C}$ . per hour to cone 14\* ( $1390^\circ \text{C}$ ). From cone 14 to cone 15 ( $1410^\circ \text{C}$ .) the temperature was raised gradually through a period of 12.5 hr., the maximum temperature reached being cone 15.5 ( $1430^\circ \text{C}$ ). The total time of firing was 60 hr. At the close of this period the gas was shut off and the furnace was allowed to cool gradually for four days, at the end of which time the temperature had fallen to  $150^\circ \text{C}$ .

On opening the furnace it was found that the brick composed of Austrian magnesite (No. 5) had retained its original form intact, but the other brick (No. 6) containing larger amounts of lime and iron oxide, had bent over, under the influence of heat and its own weight, until it just touched brick (No. 5) along its upper edge. Their appearance after firing is shown in Fig. 1. The magnesite brick (No. 5) is again shown in Fig. 2 turned at right angles to its position in Fig. 1 and showing the line along which it was touched by brick No. 6 and the area of alteration extending outward from the line. The bricks were originally eight inches high.

The Austrian magnesite brick for a distance of about one inch in every direction from the line of contact with the other brick showed a marked change in character; it had here become partially vitrified and had taken on a redder

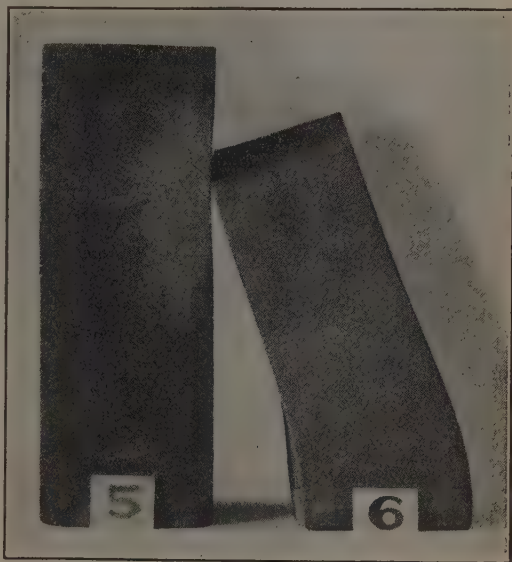


FIG. 1. *Relative positions of the bricks in the kiln; No. 5, dead-burned Austrian magnesite, No. 6, synthetic clinker.*

\*Orton pyrometric cone temperatures, as revised by the U.S. Bureau of Standards (2).

color. Between this partially vitrified area and the normal portion of the brick, was a zone darker in color, about 0.5 in. in width. This dark altered area extended into the normal brick in numerous ill-defined apophyses, feeling their way, as it were, into its substance.

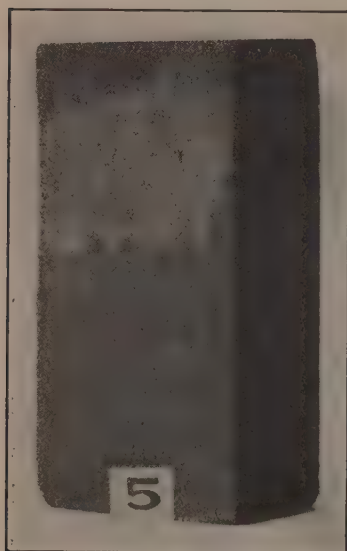


FIG. 2. View of brick No. 5 (Austrian magnesite) showing the area of alteration.

The side face of this altered brick was then divided off into a series of rectangular spaces which were numbered as shown in Fig. 3. in which the brick is shown reduced to approximately one-third of its actual size. The material enclosed in each of these areas was then analyzed—the silica, alumina, lime and iron being determined in each case, the percentage of the magnesia present being found by difference.

The results of the analysis of the material from each of these rectangular spaces and also of that of the unaltered portion of this brick (No. 5) and of the other brick (No. 6), are given in Table I.

In No. 21 and 23 the change in weight on ignition was determined and it was found that in each case there was a gain of 0.4%, indicating that a portion of the iron in the brick was present as ferrous oxide.

TABLE I  
ANALYSES OF BRICKS

No.	Percentage of			
	SiO <sub>2</sub>	Al <sub>2</sub> O <sub>3</sub>	Fe <sub>2</sub> O <sub>3</sub>	CaO
1	4.3	1.6	10.9	6.6
2	3.6	1.9	11.0	5.0
6	1.3	1.4	16.3	5.3
7	1.5	1.2	15.6	5.4
8	3.2	1.7	12.9	6.4
11	0.6	0.5	17.5	4.2
12	0.6	0.5	16.8	3.7
13	1.6	1.5	15.4	5.4
14	3.3	1.8	12.2	6.4
16	0.7	1.0	17.1	4.8
17	0.7	0.9	17.0	5.0
18	3.1	1.6	13.6	6.0
21	3.7	1.7	12.2	6.1
22	3.5	1.7	12.6	6.3
23	2.2	1.0	8.2	2.3
24	2.4	0.8	8.4	2.5
Austrian magnesite brick	2.3	0.9	8.3	2.4
Brick high in Fe <sub>2</sub> O <sub>3</sub> and CaO	1.0	1.0	22.4	8.4



These analyses show that there has been an actual transfer of lime and iron oxide (or of calcium and iron) in large amount from brick No. 6 to brick No. 5.

It is further worthy of note that this transfusion of material from one brick to another took place between two solid bodies. Brick No. 5 retained its rigidity throughout the firing and brick No. 6, while somewhat deformed was still so stiff that its edge which came against brick No. 6 was not flattened by the contact but remained sharp and clean.

The second point is that the movement took place very readily under the influence of a strong actuating force, for the whole movement of the transfused material was across the narrow knife edge where the sharp edge of one brick touched the side of the other.

A third point is that the transfusion took place quickly. While the actual time during which the two bricks were in contact is not known, it is probable that brick No. 6 did not soften sufficiently to touch brick No. 5 and start the reaction before the temperature of the furnace reached  $1310^{\circ}$  C. (cone 12), which was 45 hr. from the time the flame was turned on. If the assumption above referred to is correct, the two would have been in contact 15 hr. before the flame was turned off. If five hours be allowed as the time which elapsed until the temperature of the furnace again fell below the reaction temperature (estimate of the operator), the time in which transfusion was completed would be 20 hr.

A study of the analyses set forth in Table I shows that those of the altered portion of brick No. 5 fall into three groups, the members of each of which are approximately identical in composition but which as groups differ from one another in chemical composition in a marked manner. These several groups have been indicated in Fig. 3 by drawing a line through the rectangular spaces representing the members of each group respectively. These groups arrange themselves in three successive zones around the line of contact of the two bricks. They are as follows:—

- c. An inner zone next to the contact and consisting of divisions 11 and 12.
- b. A second zone surrounding (a) and reddish in color on the surface of this brick. This includes divisions 6, 7, 13, 16, 17.



FIG. 3. The divisions marked on the magnesite brick for purposes of analysis; the zones of like composition are marked with dotted lines.

- a. A third and outer zone, darker in color on the surface of the brick. This includes divisions 1, 8, 14, 18, 21 and 22.

The average composition of the material of these respective zones deduced from these analyses is as follows—the analyses of the two bricks by whose reactions upon one another these zones of alteration were produced are added for purposes of comparison:—

TABLE II  
COMPOSITION OF THE ALTERED ZONES

	Percentage of			
	SiO <sub>2</sub>	Al <sub>2</sub> O <sub>3</sub>	Fe <sub>2</sub> O <sub>3</sub>	CaO
Brick high in CaO and Fe <sub>2</sub> O <sub>3</sub> (No. 6)	1.0	1.0	22.4	8.4
c.	.6	.5	17.2	4.0
b.	1.6	1.3	15.6	5.4
a.	3.5	1.7	12.4	6.3
Unaltered Austrian magnesite brick (No. 5)	2.3	0.9	8.3	2.4

This table shows that the following changes have taken place:

1. Whether there has been any transfusion of SiO<sub>2</sub> from brick No. 5 to brick No. 6, is not shown but SiO<sub>2</sub> has been driven out of the highly altered portion of brick No. 5, and has accumulated in progressively greater amounts in zones b and a. In the latter zone it is present in much higher amount than it is in the unaltered brick.

2. There has been a passage of Al<sub>2</sub>O<sub>3</sub> from brick No. 6 into brick No. 5, as shown by the fact that in zone (a) the Al<sub>2</sub>O<sub>3</sub> is higher than in the unaltered rock, but otherwise all that has been said with reference to the SiO<sub>2</sub> applies to the Al<sub>2</sub>O<sub>3</sub> as well.

3. CaO has been transfused from brick No. 6 into brick No. 5, but as in the case of the SiO<sub>2</sub> and Al<sub>2</sub>O<sub>3</sub> it has been carried through the substance of the brick and been deposited in progressively larger amounts in it as the distance from the contact increases, rising from 5.4% in (b) to 6.3% in (a) beyond which the alteration ceases and the lime falls at once to 2.4% which is the lime content of the unaltered brick.

4. The Fe<sub>2</sub>O<sub>3</sub> (or Fe) has passed from brick No. 6, across the knife edge contact of the two bricks, into brick No. 5 in large amount and has been deposited in the latter in amounts which decrease progressively from the line of contact into the unaltered portion of the brick.

Thus the SiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub> and CaO in the altered portion of the brick increase progressively on receding from the line of contact from which the transfused elements came, while the amounts of Fe<sub>2</sub>O<sub>3</sub> and MgO on the other hand increase proportionately in the opposite direction, that is to say, in that part of the altered area nearest to the line of contact.

It is evident that if in the present experiment the two bricks instead of touching one another along a single sharp edge, had been placed one upon the other so that their faces were in contact and then heated, the transfusion of material from one to the other would have been much more rapid and more complete; also that out of the uniform brick of Austrian magnesite there would have been produced a brick, one side of which would be relatively rich in  $\text{SiO}_2$ ,  $\text{Al}_2\text{O}_3$  and  $\text{CaO}$  and the other relatively rich in  $\text{MgO}$  and  $\text{Fe}_2\text{O}_3$ , and this differentiation would have been brought about by heat alone.

### Discussion

The processes at work in the development of the phenomena described above are somewhat obscure. Surface tension and the development of solid solutions may play a part. Again it may be that under the influence of heat a rearrangement of the constituents in brick No. 6 took place with the development of certain more or less fluid compounds which passed into brick No. 5 and found their way through interstitial spaces into its inner parts: the brick from which it came still maintaining its essential solidity, the mass of its substance remaining intact and unaltered in character and composition. It is interesting to note if such be the case that in the altered portions of brick No. 5, the silica, alumina and lime passed outward from the line of contact of the two bricks and accumulated in continually increasing amounts toward the boundary of this altered portion, while the iron acted in exactly the opposite manner and decreased in amount as the distance from brick No. 6 increased.

Greig (3) cites a very similar instance. "The silica brick from which the roofs of open hearth steel furnaces are built take up iron oxide in use. At the same time the original lime and alumina, in the form of siliceous liquid, migrate between the grains of silica toward the cooler portions of the brick in front of an advancing liquid high in iron oxide. The brick thus becomes zoned, with a composition and structure corresponding to the original brick at the outside, followed by a zone in which there is a concentration of lime and alumina, and then by a part much higher in iron oxide. In the acid furnace this iron oxide commonly amounts to between 20 and 25 %. Tests have shown that the temperatures of melting of the parts high in iron oxides are not far below those of the original brick, and are often higher than those of the much more siliceous portion in which the alumina has concentrated."

He then shows from the equilibrium diagrams the reasons for this and states that a material, having a ferrous:ferric ratio of 1:1, in which the iron oxides form 65% of the total, the balance being silica, will be completely liquid only at  $1665^\circ\text{C}$ .

In this connexion the results of certain recent investigations are of interest.

R. B. Sosman and H. E. Merwin of the Geophysical Laboratory at Washington have investigated the system  $\text{CaO-Fe}_2\text{O}_3$  (6). A certain amount of ferrous iron is formed in all melted mixtures of  $\text{CaO}$  and  $\text{Fe}_2\text{O}_3$ . Between 50 and 100 molecular per cent  $\text{CaO}$  the percentage of ferrous oxide is small in

mixtures heated to 1400°–1500° C. They found that two compounds formed in this system on the lime side. The iron side could not be investigated. These compounds are  $2\text{CaO} \cdot \text{Fe}_2\text{O}_3$ , which melts at 1436° C. with decomposition and  $\text{CaO} \cdot \text{Fe}_2\text{O}_3$ , which melts at 1216° C., with decomposition.

More recently W. C. Hansen and L. T. Brownmiller have made equilibrium studies on alumina and ferric oxide and certain combinations with magnesia and calcium oxide (4). They claim to have identified a compound  $4\text{CaO} \cdot \text{Al}_2\text{O}_3 \cdot \text{Fe}_2\text{O}_3$  and have found evidence that this compound and  $\text{MgO}$  combine to form a series of solid solutions until the composition  $4\text{CaO} \cdot 2\text{MgO} \cdot \text{Al}_2\text{O}_3 \cdot \text{Fe}_2\text{O}_3$  is reached. Further they show facts which indicate that the solid solutions containing a greater amount of  $\text{MgO}$  than one  $\text{MgO}$  to one  $4\text{CaO} \cdot \text{Al}_2\text{O}_3 \cdot \text{Fe}_2\text{O}_3$  dissociate at  $1370^\circ \pm 5^\circ$  into liquid and  $\text{MgO}$ . ( $4\text{CaO} \cdot \text{Al}_2\text{O}_3 \cdot \text{Fe}_2\text{O}_3$  may be considered to be  $4\text{CaO} \cdot 2\text{Fe}_2\text{O}_3$ , i.e., two molecules of  $2\text{CaO} \cdot \text{Fe}_2\text{O}_3$ , in which one molecule of  $\text{Fe}_2\text{O}_3$  is replaced by one molecule of  $\text{Al}_2\text{O}_3$ .)

Furthermore evidence is deduced that a compound  $2\text{CaO} \cdot \text{MgO} \cdot \text{Fe}_2\text{O}_3$  exists and that it forms a complete series of solid solutions with  $2\text{CaO} \cdot \text{Fe}_2\text{O}_3$ . This compound and its solid solutions with  $2\text{CaO} \cdot \text{Fe}_2\text{O}_3$  dissociate on melting, forming  $\text{MgO}$  and liquid which occurs at  $1410^\circ \pm 5^\circ$ .

$\text{MgO} \cdot \text{Al}_2\text{O}_3$  and  $\text{MgO} \cdot \text{Fe}_2\text{O}_3$  form a series of solid solutions at least to a limited extent.

That the penetration of iron and lime into the Austrian brick is associated with the formation of some of the compounds and solutions above indicated would appear reasonable.

While it is inadvisable unduly to force comparisons between laboratory experiments and the great operations of nature, it is often possible to obtain from the former very valuable suggestions for the elucidation of problems presented by the latter; it would seem that the experiment under consideration suggests that certain processes may be at work in the earth's crust in zones of metamorphism which result in the production by differentiation of rock types whose origin up to the present time it has been difficult or impossible to explain. It would seem for instance that if in the geological succession a body of rock with a high content of  $\text{FeO}$ ,  $\text{Fe}_2\text{O}_3$  and  $\text{CaO}$  were to be laid down in contact with another mass high in  $\text{MgO}$ , out of the two by heat alone, without complete fusion at any rate, and without the action of water vapor or other mineralizers, an entirely new series of metamorphic rocks might be developed over an area co-extensive with the distribution of the parent rocks; this the more especially as in nature the factor of time is enormously increased, and the heat may continue to exert its action not for hours but for thousands or possibly millions of years, enabling the processes which are at work to continue in action until a perfect equilibrium is established, and the final products are developed. This apparently new factor in metamorphism may perhaps offer a solution to certain problems which have hitherto awaited a satisfactory explanation.



### **Acknowledgment**

The author desires to acknowledge his indebtedness to Messrs. F. E. Lathe, A. C. Halferdahl and J. W. Craig of the staff of the National Research Council of Canada in connexion with the preparation of this paper, and to Mr. A. F. Gill, also of the staff of the National Research Council, who made the analyses.

### **References**

1. CHAMBERLIN, T. C. *J. Geol.* 34:15. 1926.
2. FAIRCHILD, C. O. and PETERS, M. F. *J. Am. Cer. Soc.* 9:701. 1926.
3. GREIG, J. W. *Am. J. Sci.* 14:473. 1927.
4. HANSEN, W. C. and BROWNMILLER, L. T. *Am. J. Sci.* 15:225. 1928.
5. QUIRKE, T. T. *Bull. Geol. Soc. Am.* 38:763. 1927.
6. SOSMAN, R. B. and MERWIN, H. E. *J. Wash. Acad. Sci.* 6:532. 1916.

## CHROMOSOME CONDITIONS IN THE SECOND AND THIRD GENERATIONS OF PENTAPLOID WHEAT HYBRIDS<sup>1</sup>

BY J. A. JENKINS<sup>2</sup> AND W. P. THOMPSON<sup>3</sup>

### Abstract

The numbers and mating capabilities of the chromosomes were determined in a good many  $F_2$  plants and their offspring in two crosses between common (42-chromosome) wheat and emmer (28-chromosome) types. In general the results confirmed those of Kihara and supported his conclusions. Chromosome numbers were much nearer those of the parental types than was to be expected if all germ cells were capable of functioning and all zygotes of developing. A large number of expected chromosome types did not appear at all. A high percentage of  $F_2$  had only 14 bivalent chromosomes and zero to seven univalents. These tended to revert rapidly to the 14-bivalent condition of emmers.  $F_2$  in this group did not have more than 14 bivalents nor more univalents than their  $F_2$  parents. A chromosome formula for all of the group of  $F_2$  with more than 14 bivalents may be written  $(14+x)$  bivalents  $+(7-x)$  univalents. These tended to revert to the 21-bivalent condition of common wheat.  $F_2$  in this group did not have fewer bivalents nor more univalents than their  $F_2$  parents. Occasional plants were exceptions to these rules.

### Introduction

In crosses between common (42-chromosome) wheats and emmer (28-chromosome) forms it is to be expected that the parental chromosome numbers should not be found among many thousand  $F_2$ , that segregates with 29, 30, 40, or 41 should be very rare, and that ones with 34, 35, or 36 should constitute the great majority of the population. These expectations are based on the assumptions that all types of gametes are capable of functioning and all zygotes of developing. But several investigators have reported results which were quite different from these (1, 4, 6, 9, 10). Segregates which had exactly or nearly the parental numbers were frequent and those with intermediate numbers were only a small fraction of what was to be expected.

Additional important results were obtained by Kihara (1) when he determined the mating capabilities of the chromosomes in plants of the second and later generations. He found that all segregates fell naturally into two groups, (a) those in which the pollen mother cells showed 14 bivalent chromosomes (the number in the emmer parent) plus zero to seven univalents, and (b) those with more than 14 bivalents. With minor exceptions the chromosome formula of all plants in the latter group could be written  $(14+x)$  bivalents  $+(7-x)$  univalents. Since the two  $x$  cancel the sum of bivalents and univalents was 21 (the number of bivalents in the *vulgare* parent). In other words if there were more than 14 bivalents there were also enough univalents to make the total 21, that is, to give at least one complete set of the extra seven *vulgare*

<sup>1</sup> Manuscript received January 15, 1930.

Most of the work reported in this paper was done at the University of Saskatchewan, Saskatoon, Canada, with financial assistance from the National Research Council of Canada, and the rest at the Division of Genetics, University of California, Berkeley.

<sup>2</sup> Research Assistant.

<sup>3</sup> Professor of Biology, University of Saskatchewan, Saskatoon, Canada.

chromosomes. Plants with such combinations as 15 bivalents +0 to 5 univalents, 16+0 to 4, 17+0 to 3, etc., did not appear. In later generations members of group (a) tended to lose their univalents and to show only the 14-bivalent condition characteristic of pure emmers; those in group (b) gradually approached the 21-bivalent condition of pure *vulgare*. No stable type with the number of single chromosomes between 28 and 42 could be found.

The results of other investigators were not sufficiently comprehensive to test Kihara's important conclusions. Sapehin (2, 3) has indeed reported striking exceptions, for example lines with 36 chromosomes (16 bivalents+4 univalents) which bred true to that condition for several generations, and lines with constantly numerous univalents. Although Kihara's conclusions are based on the results of a great deal of work, the number of plants in  $F_2$  (20) and  $F_3$  was not large. It is possible that different kinds of material within the parental species might behave differently. Moreover further information is needed about the chromosome conditions in offspring of various chromosome types. For these reasons it seems desirable to place on record the results of a number of determinations of chromosome conditions in pentaploid hybrids, which we have made at various times, and to examine them in the light of Kihara's generalizations.

### Materials and Methods

The 42-chromosome parent was in every case the variety Marquis (*vulgare*). The parents with 28 chromosomes were White Spring Emmer (*dicoccum*) and Iumillo (*durum*).

Fixed and preserved material of most of the plants was studied in smear preparations of the pollen mother cells. In those cases in which the evidence from such preparations was insufficient for an unquestionable decision material was also studied in paraffin sections. The smears were excellent for determining the number of univalents but owing to the fact that they usually presented side views it was often impossible to determine the total number from smears alone, particularly when the number was large. In several cases therefore, the number of univalents was determined from smears and the total number from paraffin sections, but usually in determining the total the number of univalents, as previously determined, was confirmed. In cases in which the amount of material was small the examination of smears was omitted as being less likely to yield a complete determination. In a considerable number of cases it was found to be impossible to make a complete determination although we could usually be sure to within one or two chromosomes. Such counts have not been included in the results.

An effort was made to secure for study a sample representative of the whole population. It is doubtful to what extent this aim was secured because the plants on which it was impossible to make counts with complete certainty were likely to have been weak ones with few small ears.

### Experimental Data

The results of the examination of  $F_3$  plants and their  $F_2$  parents are given in Table I for *vulgare*  $\times$  *dicoccum* and in Table II for *vulgare*  $\times$  *durum*. In addition a number of exact determinations on  $F_2$  plants were made which are not included in the tables because their offspring were not studied or gave no unquestionable results. In a few cases in  $F_3$  the number of univalents only is given where this could be determined with certainty but the total number could not. In Table III the data are rearranged and summarized so as to show better the frequencies of the different combinations of bivalents and univalents and the relation of  $F_3$  to  $F_2$  results.

### Generalizations

An examination of the tables shows that all plants both in  $F_2$  and  $F_3$  had either (a) 14 bivalents +  $x$  univalents, or (b)  $(14+x)$  bivalents +  $(7-x)$  univalents, where  $x=0$  to 7. Every plant which had more than 14 bivalents had a total number (bivalents + univalents) of 21. It is therefore evident that many chromosome types which were to be expected if all types of gametes were able to function equally and all zygotes were able to develop, did not appear at all. For example an  $F_2$  with 14 bivalents plus three univalents should have produced gametes with 14 to 17 chromosomes and consequently offspring with 14, 15, 16 or 17 bivalents, 15 bivalents plus one or two univalents, and 16 bivalents plus one univalent. Actually the number of bivalents in this case was never more than 14. If it be supposed that, for some unknown reason, the chromosomes beyond 14 in number cannot mate unless a complete set of seven is present, there should have been in our example  $F_3$  with four, five or six univalents. But such plants did not occur. When the possible offspring of other  $F_2$  types are considered it becomes clear that many theoretical types never made their appearance. In general if a segregate had more than 14 bivalents it also had enough univalents to make the total of bivalents and univalents 21; there was at least one complete set of the extra seven *vulgare* chromosomes.

This general result regarding missing types and their nature agrees with the findings of Kihara. He did observe rare dwarfish "sterile combinations" with numbers approaching that of *vulgare*, namely 19 bivalents plus one univalent, and 20 bivalents. Such types were not observed in the present work.

Not only were many types missing, but the proportions of the types which did appear were not in accordance with expectations. Plants with chromosome numbers identical with, or near those of the parents were too numerous. This was particularly true in the cross *vulgare*  $\times$  *dicoccum* (see Table III).

A comparison of conditions in  $F_3$  with those in  $F_2$  shows that the descendants of plants with 14 bivalents plus zero to six univalents (Kihara's decrease group) tended to revert rapidly to the condition with 14 bivalents only, which is that of the emmer parent. With two exceptions (to be discussed later) the  $F_3$  plants in this group had no more chromosomes than their  $F_2$  parents and most



TABLE I  
CHROMOSOME CONDITIONS IN  $F_2$  AND  $F_3$  OF *T.vulgare*  $\times$  *T.dicoccum*

Second generation				Third generation			
Pedigree	Chromosomes			Pedigree	Chromosomes		
	Bivalents	Univalents	Total		Bivalents	Univalents	Total
80-3-2	14	0	14	80-3-2-2	14	0	14
				80-3-2-4	14	0	14
80-6-14	14	0	14	80-6-14-1	14	0	14
				80-6-14-2	14	0	14
				80-6-14-9	14	0	14
				80-6-14-14	14	0	14
80-3-29	14	1	15	80-3-29-2	14	1	15
80-5-16	14	1	15	80-5-16-1	14	0	14
				80-5-16-2	14	1	15
				80-5-16-6	14	0	14
				80-5-16-8	14	1	15
				80-5-16-9	14	0	14
80-5-19	14	1	15	80-5-19-1	14	0	14
				80-5-19-3	14	1	15
				80-5-19-9	14	0	14
				80-5-19-12	14	0	14
80-5-21	14	1	15	80-5-21-1	14	1	15
				80-5-21-9	14	1	15
80-2-9	14	2	16	80-2-9-3	14	1	15
				80-2-9-4	14	0	14
				80-2-9-6	14	2	16
				80-2-9-7	14	0	14
				80-2-9-11	14	0	14
				80-2-9-12	14	1	15
80-3-4	14	2	16	80-3-4-1	14	0	14
				80-3-4-2	14	0	14
				80-3-4-3	14	2	16
				80-3-4-9	14	0	14
				80-3-4-12	14	2	16
80-5-11	14	2	16	6 plants all	14	0	14
80-5-25	14	2	16	80-5-25-1	14	4	18
				80-5-25-2	14	2	16
				80-5-25-8	14	1	15
				80-5-25-10	14	0	14
				80-5-25-11	14	1	15
				80-5-25-13	14	0	14
80-2-3	14	3	17	80-2-3-1	14	0	14
				80-2-3-7	14	0	14
				80-2-3-9	14	0	14
80-2-10	14	3	17	80-2-10-5	14	2	16
				80-2-10-11		2	
				80-2-10-13	14	3	17
				80-2-10-13a	14	3	17

TABLE I—Continued

Second generation				Third generation			
Pedigree	Chromosomes			Pedigree	Chromosomes		
	Bivalents	Univalents	Total		Bivalents	Univalents	Total
80-5-10	14	3	17	80-5-10-2	14	0	14
				80-5-10-3	14	0	14
				80-5-10-4	14	1	15
				80-5-10-8		1	
				80-5-10-10	14	1	15
80-5-26	14	3	17	80-5-26-3	14	2	16
				80-5-26-5	14	0	14
				80-5-26-7	14	0	14
				80-5-26-8	14	0	14
				80-5-26-9	14	0	14
80-5-31	14	3	17	80-5-31-2	14	3	17
				80-5-31-4		2	
				80-5-31-6		0	
80-5-5	14	4	18	80-5-5-1	14	2	16
				80-5-5-3		2	
				80-5-5-4	14	0	14
				80-5-5-6		0	
				80-5-5-8	14	1	15
				80-5-5-11		1	
80-6-7	14	4	18	80-6-7-9	14	0	14
80-2-17	14	7	21	80-2-17-1	14	4	18
				80-2-17-2	14	1	15
				80-2-17-3	14	1	15
				80-2-17-4	14	3	17
				80-2-17-4a	16	5	21
				80-2-17-6	14	6	20
				80-2-17-7	18	3	21
				80-2-17-8	17	4	21
				80-2-17-9	14	6	20
				80-2-17-10	14	5	19
				80-2-17-11	14	2	16
				80-2-17-21	14	4	18
80-3-6	14	7	21	80-3-6-1	18	3	21
				80-3-6-2	17	4	21
				80-3-6-3	18	3	21
				80-3-6-7	14	2	16
				80-3-6-8	19	2	21
				80-3-6-11	14	3	17
				80-3-6-9	19	2	21
80-1-12	17	4	21	80-1-12-7	17	4	21
				80-1-12-10	17	4	21
				80-1-12-11	17	4	21
80-1-19	17	4	21	80-1-19-1	18	3	21
				80-1-19-3	17	4	21
				80-1-19-5	19	2	21
80-2-20	17	4	21	80-2-20-1	18	3	21
				80-2-20-2	18	3	21
				80-2-20-2a	18	3	21
				80-2-20-3	17	4	21
				80-2-20-5	17	4	21
				80-2-20-8	18	3	21
				80-2-20-9	17	4	21

TABLE II  
CHROMOSOME CONDITIONS IN F<sub>2</sub> AND F<sub>3</sub> OF *T. vulgare* × *T. durum*

Second generation				Third generation			
Pedigree	Chromosomes			Pedigree	Chromosomes		
	Bivalents	Univalents	Total		Bivalents	Univalents	Total
115-5-12	14	0	14	115-5-12-11	14	0	14
106-2-12	14	2	16	106-2-12-1	14	2	16
				106-2-12-2		2	
122-3-3	14	2	16	122-3-3-2	14	1	15
				122-3-3-3	14	0	14
				122-3-3-6	14	0	14
				122-3-3-9	14	7	21
				122-3-3-11	14	2	16
115-7-17	14	2	16	115-7-17-3	14	0	14
				115-7-17-6		2	
106-1-30	14	4	18	106-1-30-12	14	4	18
216-1-37	14	7	21	216-1-37-1	14	6	20
				216-1-37-2	16	5	21
				216-1-37-3	14	1	15
				216-1-37-9	19	2	21
				216-1-37-11	18	3	21
				216-1-37-12	14	2	16
				216-1-37-13	18	3	21
				216-1-37-21	16	5	21
115-1-28	16	5	21	115-1-28-1	18	3	21
				115-1-28-2	16	5	21
226-1-16	16	5	21	226-1-16-6	16	5	21
111-1-20	17	4	21	111-1-20-6		3	
				111-1-20-9	17	4	21
207-1-19	17	4	21	207-1-19-2	20	1	21
				207-1-19-3		1	
122-3-4	18	3	21	122-3-4-1	20	1	21
				122-3-4-2	19	2	21
				122-3-4-4	16	5	21
				122-3-4-7	18	3	21
				122-3-4-7a	18	3	21
				122-3-4-8	20	1	21
				122-3-4-8a	19	2	21
				122-3-4-10	21	0	21
				122-3-4-10a	20	1	21
				122-3-4-11	18	3	21
226-1-34	19	2	21	226-1-34-1	20	1	21
				226-1-34-2		1	
				226-1-34-4	19	2	21
				226-1-34-4a	19	2	21
				226-1-34-7	19	2	21
				226-1-34-9	20	1	21
				226-1-34-10	19	2	21
				226-1-34-10a	18	3	21
				226-1-34-11	20	1	21
118-2-3	20	1	21	118-2-3-1	21	0	21
				118-2-3-2	21	0	21
				118-2-3-3	20	1	21
				118-2-3-6	21	0	21
				118-2-3-8	21	0	21
				118-2-3-9	21	0	21
				118-2-3-10	21	0	21
				118-2-3-11	21	0	21

TABLE III

FREQUENCIES OF THE VARIOUS CHROMOSOME COMBINATIONS

Second generation				Third generation			
Chromosomes			Frequency	Chromosomes			Frequency
Bivalents	Univalents	Total		Bivalents	Univalents	Total	
14	0	14	3	14	0	14	7
14	1	15	4	14 14	0 1	14 15	6 6
14	2	16	7	14 14 14 14 14	0 1 2 7 4	14 15 16 21 18	16 5 5 1 1
14	3	17	5	14 14 14 14	0 1 2 3	14 15 16 17	9(10*) 2(3*) 2(4*) 3
14	4	18	3	14 14 14 14	0 1 2 4	14 15 16 18	2(3*) 1(2*) 1 1
14	7	21	4	14 14 14 14 14 14 16 17 18 19	1 2 3 4 5 6 5 4 3 2	15 16 17 18 19 20 21 21 21 21	3 3 2 2 1 3 3 2 5 3
16	5	21	2	16 18	5 3	21 21	2 1
17	4	21	5	17 18 19 20	4 3 2 1	21 21 21 21	8 5(6*) 1 1
18	3	21	2	16 18 19 20 21	5 3 2 1 0	21 21 21 21 21	1 3 2 3 1
19	2	21	1	18 19 20	3 2 1	21 21 21	1 4 3(4*)
20	1	21	1	20 21	1 0	21 21	7 1

\*Including cases in which number of univalents only was determined.



of them had fewer. The reversion in the case of  $F_2$  with 14 bivalents and two univalents seems very rapid because of one family (the offspring of 80-5-11) in which all six members showed no univalents. It is difficult to account for this case, but a re-examination of material of the  $F_2$  parent showed that the original determination was correct. There is of course a possibility of an error in the records. If this family is omitted the reversion is most rapid among the descendants with the larger numbers of univalents.

The descendants of  $F_2$  with more than 14 bivalents (Kihara's increase group) tended to revert in the other direction—towards the condition with 21 bivalents which is that of the *vulgare* parent. With two exceptions (to be discussed later) no  $F_3$  in this group had fewer bivalents than its  $F_2$  parent. Many of them had more. The reversion was not so rapid in this group as in the other. This was probably due to the loss of univalent chromosomes through lagging, for such loss would hasten the reversion in the one case and retard it in the other.

The descendants of  $F_2$  with 14 bivalents and seven univalents (the same condition as in  $F_1$ ) presented a situation similar to that in  $F_2$ , except that in this sample a larger proportion of intermediate types occurred.

Owing to the reversion in the two directions it appears that the only stable chromosome conditions were those with 14 and 21 bivalents, the conditions in the parents.

The few plants which constitute exceptions to these general rules may now be considered.  $F_3$  *vulgare*  $\times$  *durum* 122-3-3-9 (Table II) had 14 bivalents and seven univalents although its parent had only two univalents. It seems probable that this plant was the result of a natural back-pollination by a 21-chromosome male plant.  $F_3$  *vulgare*  $\times$  *dicoccum* 80-5-25-1 (Table I) had four univalents though its parent had only two. This may be a rare case in which a plant which should have had 16 bivalents was able to survive but the extra four chromosomes were unable to mate. *Vulgare*  $\times$  *durum* 122-3-4-4 and 226-1-34-10a had fewer bivalents than their parents. The explanation of their origin appears to be that one or more pairs of chromosomes in the parents had failed to mate. This phenomenon has been observed a number of times. For example a plant which undoubtedly had 19 pairs and two single chromosomes, showed occasional pollen mother cells with 18 bivalents and four univalents. Gametes resulting from such divisions might produce offspring with fewer bivalents than had their parents.

It has been mentioned that a considerable number of additional plants were studied regarding which final decisions could not be made although in most cases we could be certain to within one or two chromosomes. None of these plants were exceptions to the general rules.

The results which have been reported in this paper do not throw any new light on the reasons for the absence of many types. Inability to function normally owing to chromosome unbalance on the part of gametes with intermediate numbers may have had a big influence in determining the relative

frequency of the types which appeared. But apparently it could not account for the entire elimination of such types as those with 15 or 16 bivalents and zero to four univalents. Thompson and Cameron (8) and Sax (5) have shown that gametes with intermediate numbers do function though in a much smaller proportion of cases than is to be expected. But as long as they function at all zygotes of the types mentioned should be formed. It therefore appears necessary to attribute the result to zygotic mortality. This mortality may be due to chromosome unbalance or incompatibility or to the shrivelling of the endosperm (7).

### Acknowledgments

The authors wish to acknowledge their indebtedness to Professor E. B. Babcock for the use of the facilities of the Division of Genetics, University of California, Berkeley.

### References

1. KIHARA, H. Mem. Coll. Science, Kyoto Univ. 200 pp. 1924.
2. SAPEHIN, A. A. Bull. Applied Bot. and Plant Breeding. 19:127, 1928. (Russian with English summary).
3. SAPEHIN, A. A. and SAPEHIN, L. A. Verhand. des V internat. Kongress Vererb. 2: 1254. 1928.
4. SAX, K. Genetics 8: 301. 1923.
5. SAX, K. Verhand. des V internat. Kongress Vererb. 2: 1267. 1928.
6. THOMPSON, W. P. Genetics 10: 285. 1925.
7. THOMPSON, W. P. Genetics. 15:99. 1930.
8. THOMPSON, W. P. and Cameron, D. R. Genetics 13: 456. 1928.
9. THOMPSON, W. P. and HOLLINGSHEAD, L. J. Genetics 17: 283. 1927.
10. WATKINS, A. E. J. Genetics 15: 323. 1925.